

# Almost diagonal matrices and Besov-type spaces based on wavelet expansions \*

Markus Weimar<sup>†</sup>

August 21, 2014

## Abstract

This paper is concerned with problems in the context of the theoretical foundation of adaptive (wavelet) algorithms for the numerical treatment of operator equations. It is well-known that the analysis of such schemes naturally leads to function spaces of Besov type. But, especially when dealing with equations on non-smooth manifolds, the definition of these spaces is not straightforward. Nevertheless, motivated by applications, recently Besov-type spaces  $B_{\Psi,q}^{\alpha}(L_p(\Gamma))$  on certain two-dimensional, patchwise smooth surfaces were defined and employed successfully. In the present paper, we extend this definition (based on wavelet expansions) to a quite general class of  $d$ -dimensional manifolds and investigate some analytical properties (such as, e.g., embeddings and best  $n$ -term approximation rates) of the resulting quasi-Banach spaces. In particular, we prove that different prominent constructions of biorthogonal wavelet systems  $\Psi$  on domains or manifolds  $\Gamma$  which admit a decomposition into smooth patches actually generate the same Besov-type function spaces  $B_{\Psi,q}^{\alpha}(L_p(\Gamma))$ , provided that their univariate ingredients possess a sufficiently large order of cancellation and regularity (compared to the smoothness parameter  $\alpha$  of the space). For this purpose, a theory of almost diagonal matrices on related sequence spaces  $b_{p,q}^{\alpha}(\nabla)$  of Besov type is developed.

**Keywords:** *Besov spaces, wavelets, localization, sequence spaces, adaptive methods, non-linear approximation, manifolds, domain decomposition.*

**Subject Classification:** 35B65, 42C40, 45E99, 46A45, 46E35, 47B37, 47B38, 65T60.

---

\*This work has been supported by Deutsche Forschungsgemeinschaft DFG (DA 360/19-1).

<sup>†</sup>Philipps-University Marburg, Faculty of Mathematics and Computer Science, Hans-Meerwein-Straße, Lahnberge, 35032 Marburg, Germany. Email: weimar@mathematik.uni-marburg.de

# 1 Motivation and background

During the past years, wavelets have become a powerful tool in pure and applied mathematics. Especially for the numerical treatment of (elliptic) operator equations (which arise in models for various problems of modern sciences, e.g., in the context of image analysis, or signal processing) wavelet-based algorithms turned out to be very efficient. In these schemes wavelets are the method of choice to discretize the partial differential or integral equation under consideration such that finally an approximate solution is obtained by solving a series of finite-dimensional linear systems involving only sparsely populated matrices. Roughly speaking, this sparsity is accomplished by compression strategies which heavily exploit the multiscale structure of the wavelets on the one hand, as well as their attractive analytical properties (concerning support, cancellation, and smoothness) on the other hand. In practice, often an additional gain of performance is observed when using *adaptive* refinement and coarsening strategies that rely on (local) estimates of the residuum. Meanwhile, for many problems, this higher rate of convergence of adaptive methods (compared to non-adaptive schemes based on uniform refinement) can be justified also from the theoretical point of view. However, large problem classes remain for which a solid mathematical foundation for the need of adaptive algorithms still has to be developed.

Clearly, the best what can be expected from an adaptive wavelet solver is that it realizes (at least asymptotically) the rate of the best  $n$ -term (wavelet) approximation to the unknown solution, since this natural benchmark describes the smallest error *any* (non-linear) method can achieve using at most  $n$  degrees of freedom. When measuring the approximation error in the norm of  $L_2$ , or of Sobolev Hilbert spaces  $H^s$ , the correct smoothness spaces for calculating best  $n$ -term rates in the context of adaptive algorithms are given by (shifted versions of) the so-called *adaptivity scale of Besov spaces*  $B_\tau^\alpha(L_\tau)$ , where  $\tau^{-1} = \alpha/d + 1/2$ . Surprisingly enough, for a large class of problems including, e.g., elliptic PDEs and boundary integral equations, there exist adaptive wavelet schemes which obtain these orders of convergence, while the number of arithmetic operations stays proportional to the number of unknowns; see, e.g., [5, 6, 12, 15, 16]. In contrast, the performance of non-adaptive algorithms is governed by the maximal smoothness of the unknown solution in the Sobolev scale  $H^s$ ; cf. [8, 20]. For many practical problems this regularity is limited due to singularities caused by the shape of the underlying domain. On the other hand, particularly for elliptic PDEs on Euclidean domains, it is known that the influence of these singularities on the maximal Besov regularity is significantly smaller; see, e.g., [9, 10, 24]. Therefore we can state that, at least for such problems, adaptivity pays off and their theoretical analysis naturally leads to function spaces of Besov type.

In the realm of operator equations defined on manifolds (especially for problems formulated in terms of integral equations at the boundary of some non-smooth domain) we are

faced with additional, quite serious problems: the construction of suitable wavelet systems, on the one hand, and the investigation of the relevant function spaces, on the other hand. Meanwhile, for geometries that admit a decomposition into smooth patches (e.g., stemming from models in Computer Aided Geometric Design), a couple of wavelet bases are known which perform quite well [3, 7, 18, 19, 26, 27]. Hence, the first issue has been solved satisfactory, at least for practically important cases. The aim of this paper is to shed some light on the second problem, because the picture here is not as complete.

In classical function space theory Besov spaces are defined on the whole of  $\mathbb{R}^d$ , e.g., by Fourier techniques. Then functions (or distributions) from these spaces can be simply restricted to  $d$ -dimensional domains  $\Omega$  of interest. In all practically relevant cases this definition then coincides with intrinsic descriptions given, e.g., in terms of moduli of smoothness; cf. [33, 34, 35]. Accepting the fact that wavelet characterizations (restricted from  $\mathbb{R}^d$  to  $\Omega$ ) might involve a few wavelets whose support exceed the boundary of the domain, this method provides a handy approach towards regularity studies in Besov spaces for the case of such sets. The definition and analysis of corresponding function spaces on general manifolds is more critical: There usually trace operators or sufficiently smooth pullbacks of local (overlapping) charts are used. Unfortunately, no intrinsic characterizations for trace spaces on complicated geometries are available and it is unclear how a wavelet characterization of these spaces should look like. Moreover, following the second approach, the maximal regularity of the resulting spaces would be naturally restricted by the *global* smoothness of the manifold under consideration. Therefore, in [13] we proposed and successfully exploited a completely different method to define higher-order Besov-type spaces  $B_{\Psi,q}^\alpha(L_p(\Gamma))$  on (specific, two-dimensional, closed) manifolds  $\Gamma$  which are only *patchwise* smooth. The idea is rather simple, but quite effective: Since we like to employ wavelets for our approximation schemes, only the decay of the wavelet coefficients of the object to be approximated is important. Consequently, in our spaces we collected all those functions whose coefficients w.r.t. *one fixed* (biorthogonal) wavelet basis  $\Psi$  exhibit a certain rate of decay which would be expected from a classical wavelet characterization. Although, from the application point of view, a definition like this is completely justified, it has a theoretical drawback: The spaces constructed this way formally depend on the chosen wavelet system. In [13, Rem. 4.2(ii)] it was stated that there are good reasons to assume that spaces built up on wavelets  $\Psi$  with “similiar” properties actually coincide, at least in the sense of equivalent (quasi-)norms. The main purpose of the current article is the verification of this conjecture for a large range of parameters and three classes of wavelet bases ( $\Psi_{\text{DS}}$ ,  $\Psi_{\text{HS}}$ , and  $\Psi_{\text{CTU}}$ ) which are widely used in practice; see Theorem 4.3.

Our material is organized as follows: Section 2 exclusively deals with (operators on) sequence spaces  $b_{p,q}^\alpha(\nabla)$  which later on will be crucial for the definition of our function spaces  $B_{\Psi,q}^\alpha(L_p(\Gamma))$  of Besov-type. They are indexed by what we call *multiscale grids*  $\nabla$ , i.e., by

sets which are tailored for the use in the context of multiresolution expansions on manifolds. Furthermore, guided by the pioneer work of Frazier and Jawerth [23], here we introduce classes  $\text{ad}(b_{p,q}^{\alpha_0}(\nabla^0), b_{p,q}^{\alpha_1}(\nabla^1))$  of *almost diagonal* matrices whose entries decay sufficiently fast apart from the diagonal. The main result of this first part (which is of interest on its own, but also will be essential in what follows) then is given by Theorem 2.9. It states that every such matrix induces a bounded linear operator between the Besov-type sequence spaces  $b_{p,q}^{\alpha_i}(\nabla^i)$ ,  $i = 0, 1$ , under consideration. The remaining part of the paper is concerned with function spaces. In Section 3 we recall what is meant by patchwise smooth geometries  $\Gamma$  and we review basic concepts from multiscale analysis. Additionally, here we describe the three biorthogonal wavelet constructions  $\Psi = (\Psi^\Gamma, \tilde{\Psi}^\Gamma)$  we are mainly interested in:

- 1.) Composite wavelet bases  $\Psi = \Psi_{\text{DS}}$  introduced by Dahmen and Schneider [18] (for general operator equations),
- 2.) Modified composite wavelets  $\Psi = \Psi_{\text{HS}}$  due to Harbrecht and Stevenson [27] (which are the first choice in the so-called boundary element method for integral equations), and
- 3.) Bases  $\Psi = \Psi_{\text{CTU}}$  (primarily used in the wavelet element method) established by Canuto, Tabacco, Urban [2, 3].

Afterwards, in Section 3.3, we extend the definition of Besov-type function spaces  $B_{\Psi,q}^\alpha(L_p(\Gamma))$  given in [13] to a fairly general setting which particularly covers spaces on decomposable manifolds  $\Gamma$  that are needed for practical applications. Moreover, here we investigate some of the theoretical properties of these scales such as embeddings, interpolation results, and best  $n$ -term approximation rates. In Theorem 4.2, Section 4, we employ the theory of almost diagonal matrices developed in Section 2, to derive sufficient conditions for continuous one-sided *change of basis embeddings*

$$B_{\Psi,q}^\alpha(L_p(\Gamma)) \hookrightarrow B_{\Phi,q}^\alpha(L_p(\Gamma)), \quad 0 \leq \alpha < \alpha^*.$$

Finally, these embeddings then pave the way to state and prove Theorem 4.3 which constitutes the main result of this paper: the equivalence  $B_{\Psi,q}^\alpha(L_p(\Gamma)) = B_{\Phi,q}^\alpha(L_p(\Gamma))$  for  $\Psi, \Phi \in \{\Psi_{\text{DS}}, \Psi_{\text{HS}}, \Psi_{\text{CTU}}\}$ . The article is concluded with an appendix (Section 5) which contains auxiliary assertions, as well as some quite technical proofs.

**Notation:** For families  $\{a_{\mathcal{J}}\}_{\mathcal{J}}$  and  $\{b_{\mathcal{J}}\}_{\mathcal{J}}$  of non-negative real numbers over a common index set we write  $a_{\mathcal{J}} \lesssim b_{\mathcal{J}}$  if there exists a constant  $c > 0$  (independent of the context-dependent parameters  $\mathcal{J}$ ) such that

$$a_{\mathcal{J}} \leq c \cdot b_{\mathcal{J}}$$

holds uniformly in  $\mathcal{J}$ . Consequently,  $a_{\mathcal{J}} \sim b_{\mathcal{J}}$  means  $a_{\mathcal{J}} \lesssim b_{\mathcal{J}}$  and  $b_{\mathcal{J}} \lesssim a_{\mathcal{J}}$ . In addition, if not further specified, throughout the whole paper we will assume that  $\Gamma$  denotes an arbitrary

set endowed with some metric  $\varrho_\Gamma$ . In view of the applications we have in mind, later on we will impose additional conditions on  $\Gamma$ .

## 2 Sequence spaces and almost diagonal matrices

For every  $\mathcal{T} \neq \emptyset$  let us define a *pseudometric* on  $\Gamma \times \mathcal{T}$  by setting

$$\text{dist}((y, t), (y', t')) := \varrho_\Gamma(y, y') \quad \text{for } y, y' \in \Gamma \quad \text{and } t, t' \in \mathcal{T}. \quad (1)$$

**Definition 2.1.** *Let  $d \in \mathbb{N}$ . We say the sequence  $\nabla := (\nabla_j)_{j \in \mathbb{N}_0}$  is a multiscale grid of dimension  $d$  (for  $\Gamma$ ) if there are absolute constants  $c_1, c_2, c_3 > 0$  such that the following three assumptions are satisfied:*

(A1) *For some finite index set  $\mathcal{T} \neq \emptyset$  and all  $j \in \mathbb{N}_0$  the set  $\nabla_j$  forms a  $(c_1 2^{-j})$ -net for  $\Gamma \times \mathcal{T}$  w.r.t. (1).*

(A2)  *$\nabla$  is uniformly well-separated, meaning that uniformly in  $j \in \mathbb{N}_0$  it holds*

$$\sup_{\xi \in \nabla_j} \#\{\xi' \in \nabla_j \mid \text{dist}(\xi, \xi') \leq c_2 2^{-j}\} \lesssim 1.$$

(A3)  *$\nabla$  is uniformly  $d$ -dimensional, in the sense that uniformly in  $j \in \mathbb{N}_0$  we have*

$$\sup_{\xi \in \nabla_j} \#\{\xi' \in \nabla_j \mid \text{dist}(\xi, \xi') \leq c_3\} \sim 2^{dj}.$$

**Remark 2.2.** We note in passing that these assumptions clearly force  $\Gamma$  to be  $d$ -dimensional (in a certain sense), as well. Moreover, if the set  $\Gamma$  is bounded, meaning that its diameter  $\text{diam}(\Gamma) := \sup_{y, y' \in \Gamma} \varrho_\Gamma(y, y')$  is finite, then Definition 2.1 implies that

(A4a)  $\#\nabla_j \sim 2^{dj}$ .

Otherwise (if  $\Gamma$  is unbounded), we necessarily have

(A4b)  $\#\nabla_j = \infty$  for all  $j \in \mathbb{N}_0$ . □

Typical examples of multiscale grids cover index sets related to expansions (w.r.t. certain building blocks such as wavelets, atoms, molecules, ...) in function spaces on  $\Gamma = \mathbb{R}^d$ ,  $d \in \mathbb{N}$ . Indeed, when dealing with wavelet expansions,  $\tilde{\nabla}_j \subseteq \mathbb{Z}^d \times \{1, \dots, 2^d - 1\}$  usually is interpreted as index set encoding the location and type of all wavelets at level  $j \in \mathbb{N}_0$ . The same reasoning also applies for (bounded) domains  $\Gamma \subsetneq \mathbb{R}^d$ . Obviously, every such sequence  $\tilde{\nabla} = (\tilde{\nabla}_j)_{j \in \mathbb{N}_0}$  can be identified with some multiscale grid  $\nabla$  in the above sense. However, note that the index sets in Definition 2.1 are designed in a way such that all indices are

directly associated with some point *in the domain*  $\Gamma$  which allows to handle more complex situations, as well. If  $d = 2$ , say, then we may also think of  $\Gamma = \partial\Omega$  being the surface of some bounded polyhedral domain  $\Omega \subset \mathbb{R}^3$ , or an even more abstract (two-dimensional) manifold with or without boundary. Then for all  $j \in \mathbb{N}_0$  the sets

$$\{y \in \Gamma \mid (y, t) = \xi \in \nabla_j \text{ for some } t \in \mathcal{T}\}$$

yield discretizations of  $\Gamma$ . Let us mention that for infinitely smooth manifolds  $\Gamma$  a similar approach has been proposed already in [34].

Now we are well-prepared to introduce *sequence spaces*  $b_{p,q}^\alpha(\nabla)$  of *Besov type* on multiscale grids  $\nabla$ .

## 2.1 Sequence spaces of Besov type

The following Definition 2.3 is inspired by sequence spaces which naturally arise in the context of classical (wavelet) characterizations of function spaces; see, e.g., [11, Def. 3]. Here and in what follows we slightly abuse the notation and write  $(j, \xi) \in \nabla = (\nabla_j)_{j \in \mathbb{N}_0}$  if  $j \in \mathbb{N}_0$  and  $\xi \in \nabla_j$ .

**Definition 2.3.** Let  $0 < p < \infty$ ,  $0 < q \leq \infty$ , as well as  $\alpha \in \mathbb{R}$ , and let  $\nabla = (\nabla_j)_{j \in \mathbb{N}_0}$  denote some multiscale grid of dimension  $d \in \mathbb{N}$  in the sense of Definition 2.1. Then

- (i) we define the sequence space  $b_{p,q}^\alpha(\nabla) := \{\mathbf{a} = (a_{(j,\xi)})_{(j,\xi) \in \nabla} \subset \mathbb{C} \mid \|\mathbf{a}\|_{b_{p,q}^\alpha(\nabla)} < \infty\}$  endowed with the (quasi-)norm

$$\|\mathbf{a}\|_{b_{p,q}^\alpha(\nabla)} := \begin{cases} \left( \left( \sum_{j=0}^{\infty} 2^{j(\alpha+d[1/2-1/p])} \right)^q \left[ \sum_{\xi \in \nabla_j} |a_{(j,\xi)}|^p \right]^{q/p} \right)^{1/q}, & \text{if } q < \infty, \\ \sup_{j \in \mathbb{N}_0} 2^{j(\alpha+d[1/2-1/p])} \left[ \sum_{\xi \in \nabla_j} |a_{(j,\xi)}|^p \right]^{1/p}, & \text{if } q = \infty. \end{cases} \quad (2)$$

- (ii) we set  $\sigma_p := \sigma_p(d) := d \cdot \max\left\{\frac{1}{p} - 1, 0\right\}$ .

**Remark 2.4.** Some comments are in order:

- (i) Using standard arguments it can be checked that  $b_{p,q}^\alpha(\nabla)$  are always quasi-Banach spaces. They are Banach spaces if and only if  $\min\{p, q\} \geq 1$ , and Hilbert spaces if and only if  $p = q = 2$ .

- (ii) For special choices of the parameters  $p$ ,  $q$ , and  $\alpha$ , the (quasi-)norms defined in (2) simplify significantly. In particular,

$$\|\mathbf{a}\|_{b_{2,2}^0(\nabla)} = \left( \sum_{(j,\xi) \in \nabla} |a_{(j,\xi)}|^2 \right)^{1/2},$$

such that  $b_{2,2}^0(\nabla) = \ell_2(\nabla)$  with equal norms. More general we have  $b_{\tau,\tau}^{\alpha_\tau}(\nabla) = \ell_\tau(\nabla)$  with equal norms, where

$$\tau = \left( \frac{\alpha_\tau}{d} + \frac{1}{2} \right)^{-1} \quad \text{with} \quad \alpha_\tau \geq 0 \quad (3)$$

defines the so-called *adaptivity scale* w.r.t.  $\ell_2(\nabla)$ . □

Before we turn to operators acting on the sequence spaces just introduced, let us add the following embedding result which will be useful later on. Its proof is postponed to the appendix; see Section 5.2.

**Proposition 2.5** (Standard embeddings). *Let  $0 < p_0, p_1 < \infty$ , as well as  $0 < q_0, q_1 \leq \infty$ , and  $\alpha, \gamma \in \mathbb{R}$ . Moreover, let  $\nabla$  denote some multiscale grid of dimension  $d \in \mathbb{N}$ . If, in addition, condition (A4a) holds for  $\nabla$  then the embedding*

$$b_{p_0, q_0}^{\alpha+\gamma}(\nabla) \hookrightarrow b_{p_1, q_1}^{\alpha}(\nabla)$$

*exists (as a set theoretic inclusion) if and only if it is continuous if and only if one of the subsequent conditions applies:*

- )  $\gamma > d \cdot \max\left\{0, \frac{1}{p_0} - \frac{1}{p_1}\right\},$
- )  $\gamma = d \cdot \max\left\{0, \frac{1}{p_0} - \frac{1}{p_1}\right\} \quad \text{and} \quad q_0 \leq q_1.$

*Furthermore, if  $\nabla$  satisfies (A4b) instead of (A4a) then a corresponding characterization holds with the additional condition  $p_0 \leq p_1$ .*

**Remark 2.6.** Based on the reduction arguments we used to prove Proposition 2.5 it would be possible to derive a lot of further results related to the spaces  $b_{p,q}^{\alpha}(\nabla)$  such as, e.g., interpolation assertions or estimates for entropy numbers. As such properties are beyond the scope of the present paper, we will not follow this line of research here. □

## 2.2 Almost diagonal matrices

Clearly, every linear mapping  $M$  defined between sequence spaces (indexed by multiscale grids  $\nabla^0$  and  $\nabla^1$  on  $\Gamma$ , respectively) can be represented as the formal product with some double-infinite complex matrix

$$\mathcal{M} := \{m_{(j,\xi),(k,\eta)}\}_{(j,\xi) \in \nabla^1, (k,\eta) \in \nabla^0},$$

i.e.,  $M: \mathbf{a} \mapsto M\mathbf{a} := ((\mathcal{M}\mathbf{a})_{(j,\xi)})_{(j,\xi) \in \nabla^1}$  with

$$(\mathcal{M}\mathbf{a})_{(j,\xi)} := \sum_{(k,\eta) \in \nabla^0} m_{(j,\xi),(k,\eta)} a_{(k,\eta)}, \quad (j,\xi) \in \nabla^1.$$

We shall follow the ideas given in [23, Sect. 3] and define classes of *almost diagonal* matrices for the sequence spaces of Besov type established in Definition 2.3.

**Definition 2.7.** Let  $0 < p < \infty$  and  $0 < q \leq \infty$ . Moreover, let  $\nabla^0$  and  $\nabla^1$  denote two multiscale grids of dimension  $d \in \mathbb{N}$  for some set  $\Gamma$ .

(i) For  $\alpha_0, \alpha_1 \in \mathbb{R}$  a matrix  $\mathcal{M} = \{m_{(j,\xi),(k,\eta)}\}_{(j,\xi) \in \nabla^1, (k,\eta) \in \nabla^0}$  is called *almost diagonal* between  $b_{p,q}^{\alpha_0}(\nabla^0)$  and  $b_{p,q}^{\alpha_1}(\nabla^1)$  if there exists  $\varepsilon > 0$  such that

$$\sup_{(j,\xi) \in \nabla^1, (k,\eta) \in \nabla^0} \frac{|m_{(j,\xi),(k,\eta)}|}{\omega_{(j,\xi),(k,\eta)}(\varepsilon)} < \infty, \quad (4)$$

where

$$\omega_{(j,\xi),(k,\eta)}(\varepsilon) := 2^{k\alpha_0 - j\alpha_1} \cdot \frac{\min\{2^{-(j-k)(d/2+\varepsilon)}, 2^{(j-k)(d/2+\varepsilon+\sigma_p)}\}}{[1 + \min\{2^k, 2^j\} \text{dist}(\xi, \eta)]^{d+\varepsilon+\sigma_p}}.$$

In this case we write  $\mathcal{M} \in \text{ad}(b_{p,q}^{\alpha_0}(\nabla^0), b_{p,q}^{\alpha_1}(\nabla^1))$ .

(ii) When there is no danger of confusion, we shall write  $\text{ad}_p^{\alpha_0, \alpha_1}$  as a shortcut for the class  $\text{ad}(b_{p,q}^{\alpha_0}(\nabla^0), b_{p,q}^{\alpha_1}(\nabla^1))$ . Furthermore, if  $\alpha_0 = \alpha_1 = \alpha \in \mathbb{R}$  then we set  $\text{ad}_p^\alpha := \text{ad}_p^{\alpha_0, \alpha_1}$ .

Roughly speaking, a matrix  $\mathcal{M}$  belongs to the class  $\text{ad}_p^{\alpha_0, \alpha_1}$  if its entries decay fast enough apart from the diagonal  $(m_{(j,\xi),(j,\xi)})$ . If the sets  $\nabla_j^i$ ,  $i \in \{0, 1\}$ , are interpreted as index sets for the location (and type) of all wavelets at level  $j$  on  $\Gamma$  then the matrix entries  $m_{(j,\xi),(k,\eta)}$  need to be small for all wavelets (indexed by  $(j, \xi) \in \nabla^1$  and  $(k, \eta) \in \nabla^0$ , respectively) which

- are supported far away from each other (then  $\text{dist}(\xi, \eta) \gg 0$ ), or
- correspond to very different levels (then  $|j - k| \gg 0$ ).



Note that quite similar matrix classes naturally appear in the context of compression and preconditioning strategies used by elaborated (adaptive) wavelet algorithms for operator equations (e.g., for Schwartz kernel problems). Without going into details, we like to mention the so-called *Lemarié algebra* and refer to [4, 5, 6, 14, 29] for details.

**Remark 2.8.** Some further comments are in order:

- (i) Observe that (4) is independent of the index  $q$  which justifies to suppress this fine-tuning parameter in the abbreviations in Definition 2.7(ii).
- (ii) Using the monotonicity of  $\sigma_p$  (cf. Definition 2.3(ii)), it is easily seen that

$$\text{ad}_p^{\alpha_0, \alpha_1} \subseteq \text{ad}_p^{\alpha_0, \alpha_1} \subseteq \text{ad}_1^{\alpha_0, \alpha_1} = \text{ad}_p^{\alpha_0, \alpha_1} \quad \text{for all} \quad 0 < \hat{p} \leq p \leq 1 \leq \tilde{p} < \infty$$

and all  $\alpha_0, \alpha_1 \in \mathbb{R}$ . That is, condition (4) is getting stronger when  $1/p$  increases and it is independent of  $p$  when  $1/p \leq 1$ .  $\square$

We are ready to state and prove the main result of this Section 2. It is inspired by [23, Thm. 3.3] and shows that every almost diagonal matrix can be interpreted as a continuous linear operator on the sequence spaces introduced in Definition 2.3.

**Theorem 2.9.** *Let  $0 < p < \infty$  and  $0 < q \leq \infty$ , as well as  $\alpha_0, \alpha_1 \in \mathbb{R}$ . Moreover, let  $\nabla^0$  and  $\nabla^1$  denote two multiscale grids of dimension  $d \in \mathbb{N}$  for some set  $\Gamma$ . Then every matrix  $\mathcal{M} \in \text{ad}(b_{p,q}^{\alpha_0}(\nabla^0), b_{p,q}^{\alpha_1}(\nabla^1))$  induces a bounded linear operator  $M: b_{p,q}^{\alpha_0}(\nabla^0) \rightarrow b_{p,q}^{\alpha_1}(\nabla^1)$ .*

Before we conclude this section by presenting a detailed derivation of this assertion, let us remark that Theorem 2.9 (as well as its proof) differs from [23, Thm. 3.3] to some extent. Indeed, in the current paper, we deal with sequence spaces which correspond to function spaces of Besov type, in contrast to Triebel-Lizorkin spaces discussed in [23]. Consequently, as we shall see, we can avoid the use of maximal inequalities due to Fefferman-Stein. Another difference is that our notion of almost diagonal matrices (and thus also Theorem 2.9) depends on two smoothness indices  $\alpha_0$  and  $\alpha_1$  which might be useful in applications. Moreover, the authors of [23] needed to apply duality results in their proof (to handle the term which corresponds to  $\mathcal{M}^+$  in Step 2 of our proof given below). This is not necessary in our case, but it would be possible of course. Finally, due to the mild assumptions on the two (possibly different) multiscale grids  $\nabla^i$ ,  $i \in \{0, 1\}$ , in sharp contrast to [23], our theorem is not restricted to spaces related to function spaces on the whole of  $\mathbb{R}^d$ . Indeed, in Section 4, we will employ Theorem 2.9 to derive a result for Besov-type function spaces on (bounded) manifolds or domains, respectively.

*Proof (of Theorem 2.9).* Let  $0 < q \leq \infty$ . Following the lines of [23], we split the proof into three parts corresponding to different parameter constellations. To keep the presentation

as streamlined as possible we moreover postpone some technicalities to the appendix in Section 5.1.

*Step 1* (case  $\alpha_0 = \alpha_1 = 0$  and  $1 < p < \infty$ ). For  $p > 1$  let  $\mathcal{M} \in \text{ad}_p^0$  and  $\mathbf{a} \in b_{p,q}^0(\nabla^0)$ . Writing  $\mathcal{M} = \mathcal{M}^- + \mathcal{M}^+$ , where we set

$$(\mathcal{M}^- \mathbf{a})_{(j,\xi)} := \sum_{0 \leq k < j} \sum_{\eta \in \nabla_k^0} m_{(j,\xi),(k,\eta)} a_{(k,\eta)} \quad \text{and} \quad (\mathcal{M}^+ \mathbf{a})_{(j,\xi)} := \sum_{k \geq j} \sum_{\eta \in \nabla_k^0} m_{(j,\xi),(k,\eta)} a_{(k,\eta)}$$

for every  $(j, \xi) \in \nabla^1$ , we have to show that the associated linear operators  $M^-$  and  $M^+$  are bounded mappings from  $b_{p,q}^0(\nabla^0)$  into itself, i.e., that

$$\|M^\pm \mathbf{a}\|_{b_{p,q}^0(\nabla^1)} \leq c \|\mathbf{a}\|_{b_{p,q}^0(\nabla^0)} \quad (5)$$

with some  $c > 0$  independent of  $\mathbf{a}$ .

Let us first consider  $\mathcal{M}^-$  and  $M^-$ , respectively. Since  $1 < p < \infty$ , the triangle inequality together with Minkowski's inequality yields for every fixed  $j \in \mathbb{N}_0$ ,

$$\begin{aligned} \left[ \sum_{\xi \in \nabla_j^1} |(\mathcal{M}^- \mathbf{a})_{(j,\xi)}|^p \right]^{1/p} &\leq \left[ \sum_{\xi \in \nabla_j^1} \left( \sum_{0 \leq k < j} \left| \sum_{\eta \in \nabla_k^0} m_{(j,\xi),(k,\eta)} a_{(k,\eta)} \right| \right)^p \right]^{1/p} \\ &\leq \sum_{0 \leq k < j} \left( \sum_{\xi \in \nabla_j^1} \left| \sum_{\eta \in \nabla_k^0} m_{(j,\xi),(k,\eta)} a_{(k,\eta)} \right|^p \right)^{1/p}. \end{aligned}$$

Due to Definition 2.7 we have  $|m_{(j,\xi),(k,\eta)}| \leq C \cdot \omega_{(j,\xi),(k,\eta)}(\varepsilon)$  for some  $C = C(\varepsilon) > 0$  and all  $(j, \xi) \in \nabla^1$ ,  $(k, \eta) \in \nabla^0$ . Note that  $\sigma_p = 0$ , since  $1 < p < \infty$ . Thus, we conclude

$$\left[ \sum_{\xi \in \nabla_j^1} |(\mathcal{M}^- \mathbf{a})_{(j,\xi)}|^p \right]^{1/p} \lesssim \sum_{0 \leq k < j} 2^{(k-j)(d/2+\varepsilon)} \left( \sum_{\xi \in \nabla_j^1} \left| \sum_{\eta \in \nabla_k^0} \frac{|a_{(k,\eta)}|}{[1 + 2^k \text{dist}(\xi, \eta)]^{d+\varepsilon}} \right|^p \right)^{1/p} \quad (6)$$

for every  $j \in \mathbb{N}_0$ . Observe that for every fixed  $I \in \mathbb{N}_0$  the sets  $\nabla_I^i$ ,  $i \in \{0, 1\}$ , equipped with the counting measure  $\mu_I$  form  $\sigma$ -finite measure spaces and that  $L_p(\nabla_I^i, \mu_I) = \ell_p(\nabla_I^i)$ . Hence, for every  $j, k \in \mathbb{N}_0$  with  $0 \leq k < j$ , we can rewrite the sum in the brackets as

$$\left\| T_{j,k,\varepsilon}^- (|a_{(k,\eta)}|)_{\eta \in \nabla_k^0} \right\|_{\ell_p(\nabla_j^1)},$$

where  $T_{j,k,\varepsilon}^- : \ell_p(\nabla_k^0) \rightarrow \ell_p(\nabla_j^1)$  is an integral (summation) operator with kernel

$$K_{j,k,\varepsilon}^-(\xi, \eta) := \frac{1}{[1 + 2^k \text{dist}(\xi, \eta)]^{d+\varepsilon}}, \quad \xi \in \nabla_j^1, \eta \in \nabla_k^0.$$

Using a Schur-type argument (see Lemma 5.2) together with Lemma 5.3 from Section 5.1 below, for  $s = d + \varepsilon > d$  its operator norm can be bounded by

$$\begin{aligned} \|T_{j,k,\varepsilon}^- \big| \mathcal{L}(\ell_p(\nabla_k^0), \ell_p(\nabla_j^1))\| &\leq \left( \sup_{\eta \in \nabla_k^0} \sum_{\xi \in \nabla_j^1} |K_{j,k,\varepsilon}^-(\xi, \eta)| \right)^{1/p} \left( \sup_{\xi \in \nabla_j^1} \sum_{\eta \in \nabla_k^0} |K_{j,k,\varepsilon}^-(\xi, \eta)| \right)^{1/p'} \\ &\leq C' (\max\{1, 2^{(j-k)(d+\varepsilon)}\})^{1/p} (\max\{1, 2^{(k-j)(d+\varepsilon)}\})^{1/p'} \\ &= C' 2^{(j-k)(d+\varepsilon)/p}, \end{aligned} \quad (7)$$

where  $1/p' + 1/p = 1$  and  $C' > 0$  does not depend on  $j$  and  $k$ . Hence, from (6) it follows

$$\begin{aligned} \left[ \sum_{\xi \in \nabla_j^1} |(\mathcal{M}^- \mathbf{a})_{(j,\xi)}|^p \right]^{1/p} &\lesssim \sum_{0 \leq k < j} 2^{(k-j)(d/2+\varepsilon)} 2^{(j-k)(d+\varepsilon)/p} \left( \sum_{\eta \in \nabla_k^0} |a_{(k,\eta)}|^p \right)^{1/p} \\ &= 2^{-jd[\frac{1}{2}-\frac{1}{p}]} \sum_{0 \leq k < j} 2^{-(j-k)\varepsilon/p'} \left[ 2^{kd[\frac{1}{2}-\frac{1}{p}]} \left( \sum_{\eta \in \nabla_k^0} |a_{(k,\eta)}|^p \right)^{1/p} \right]. \end{aligned} \quad (8)$$

Finally, we multiply by  $2^{jd[\frac{1}{2}-\frac{1}{p}]}$ , take the  $\ell_q$ -(quasi-)norm with respect to  $j \in \mathbb{N}_0$ , and apply Lemma 5.1 (with  $\delta := \varepsilon/p' > 0$  and  $r := 1$ ) to obtain

$$\left\{ \sum_{j \in \mathbb{N}_0} 2^{jd[\frac{1}{2}-\frac{1}{p}]q} \left[ \sum_{\xi \in \nabla_j^1} |(\mathcal{M}^- \mathbf{a})_{(j,\xi)}|^p \right]^{q/p} \right\}^{1/q} \lesssim \left\{ \sum_{k \in \mathbb{N}_0} 2^{kd[\frac{1}{2}-\frac{1}{p}]q} \left[ \sum_{\eta \in \nabla_k^0} |a_{(k,\eta)}|^p \right]^{q/p} \right\}^{1/q},$$

if  $q < \infty$ , and

$$\sup_{j \in \mathbb{N}_0} 2^{jd[\frac{1}{2}-\frac{1}{p}]} \left[ \sum_{\xi \in \nabla_j^1} |(\mathcal{M}^- \mathbf{a})_{(j,\xi)}|^p \right]^{1/p} \lesssim \sup_{k \in \mathbb{N}_0} 2^{kd[\frac{1}{2}-\frac{1}{p}]} \left[ \sum_{\eta \in \nabla_k^0} |a_{(k,\eta)}|^p \right]^{1/p},$$

if  $q = \infty$ , respectively. Hence, we have shown (5) for  $M^-$ .

We turn to  $\mathcal{M}^+$  and  $M^+$ , respectively. The analogue of (6) for fixed  $j \in \mathbb{N}_0$  reads

$$\left[ \sum_{\xi \in \nabla_j^1} |(\mathcal{M}^+ \mathbf{a})_{(j,\xi)}|^p \right]^{1/p} \lesssim \sum_{k \geq j} 2^{(j-k)(d/2+\varepsilon)} \left( \sum_{\xi \in \nabla_j^1} \left| \sum_{\eta \in \nabla_k^0} \frac{|a_{(k,\eta)}|}{[1 + 2^j \text{dist}(\xi, \eta)]^{d+\varepsilon}} \right|^p \right)^{1/p}$$

such that the kernel of the associated integral operator  $T_{j,k,\varepsilon}^+ : \ell_p(\nabla_k^0) \rightarrow \ell_p(\nabla_j^1)$  is given by

$$K_{j,k,\varepsilon}^+(\xi, \eta) := \frac{1}{[1 + 2^j \text{dist}(\xi, \eta)]^{d+\varepsilon}}, \quad \xi \in \nabla_j^1, \eta \in \nabla_k^0.$$

Hence, (7) is replaced by  $\|T_{j,k,\varepsilon}^+\| \leq C' 2^{(k-j)(d+\varepsilon)/p'}$  such that (8) now reads

$$\left[ \sum_{\xi \in \nabla_j^1} |(\mathcal{M}^+ \mathbf{a})_{(j,\xi)}|^p \right]^{1/p} \lesssim 2^{-jd[\frac{1}{2}-\frac{1}{p}]} \sum_{k \geq j} 2^{(j-k)\varepsilon/p} \left[ 2^{kd[\frac{1}{2}-\frac{1}{p}]} \left( \sum_{\eta \in \nabla_k^0} |a_{(k,\eta)}|^p \right)^{1/p} \right].$$

Since  $\delta := \varepsilon/p > 0$ , the assertion thus follows as before. This shows (5) also for  $M^+$  and hence it completes Step 1.

*Step 2 (case  $\alpha_0 = \alpha_1 = 0$  and  $0 < p \leq 1$ ).* Let  $\mathbf{a} = (a_{(k,\eta)})_{(k,\eta) \in \nabla^0} \in b_{p,q}^0(\nabla^0)$  and choose  $0 < r < p \leq 1$ , i.e.,  $1 < p/r < \infty$ . For every such  $r$  we define  $\tilde{\mathbf{a}} := \tilde{\mathbf{a}}(r) := (\tilde{a}_{(k,\eta)})_{(k,\eta) \in \nabla^0}$  by  $\tilde{a}_{(k,\eta)} := 2^{-kd[\frac{1}{2}-\frac{r}{2}]} |a_{(k,\eta)}|^r$ ,  $(k, \eta) \in \nabla^0$ , so that  $\|\mathbf{a}\|_{b_{p,q}^0(\nabla^0)} = \|\tilde{\mathbf{a}}\|_{b_{p/r,q/r}^0(\nabla^0)}^{1/r}$ , i.e.,  $\tilde{\mathbf{a}} \in b_{p/r,q/r}^0(\nabla^0)$ . Similarly, given a matrix  $\mathcal{M} = \{m_{(j,\xi),(k,\eta)}\}_{(j,\xi) \in \nabla^1, (k,\eta) \in \nabla^0}$ , we set

$$\tilde{\mathcal{M}} := \tilde{\mathcal{M}}(r) := \{\tilde{m}_{(j,\xi),(k,\eta)}\}_{(j,\xi) \in \nabla^1, (k,\eta) \in \nabla^0} := \left\{ 2^{(k-j)d[\frac{1}{2}-\frac{r}{2}]} |m_{(j,\xi),(k,\eta)}|^r \right\}_{(j,\xi) \in \nabla^1, (k,\eta) \in \nabla^0}.$$

If we assume that there exists  $\varepsilon > 0$  such that  $\mathcal{M}$  belongs to  $\text{ad}_p^0$  then straightforward calculations show that  $\tilde{\mathcal{M}} = \tilde{\mathcal{M}}(r) \in \text{ad}_{p/r}^0$  with  $\tilde{\varepsilon} := \varepsilon r + d(r/p - 1)$ . Note that  $\tilde{\varepsilon} > 0$ , provided that we restrict ourselves to  $r$  with  $pd/(\varepsilon p + d) < r < p$  (which can be done w.l.o.g.). Using Jensen's inequality we obtain

$$\left| (\mathcal{M} \mathbf{a})_{(j,\xi)} \right|^p \leq \left( \sum_{(k,\eta) \in \nabla^0} |m_{(j,\xi),(k,\eta)} a_{(k,\eta)}|^r \right)^{p/r} = 2^{jd[\frac{1}{2}-\frac{r}{2}]p/r} \left| (\tilde{\mathcal{M}} \tilde{\mathbf{a}})_{(j,\xi)} \right|^{p/r}$$

for every  $(j, \xi) \in \nabla^1$ . Therefore, for  $q < \infty$ , the associated operator  $M$  satisfies

$$\begin{aligned} \|M \mathbf{a}\|_{b_{p,q}^0(\nabla^1)}^r &= \left( \sum_{j \in \mathbb{N}_0} 2^{jd[\frac{1}{2}-\frac{1}{p}]q} \left[ \sum_{\xi \in \nabla_j^1} |(\mathcal{M} \mathbf{a})_{(j,\xi)}|^p \right]^{q/p} \right)^{r/q} \\ &\leq \left( \sum_{j \in \mathbb{N}_0} 2^{jd[\frac{1}{2}-\frac{1}{p/r}]q/r} \left[ \sum_{\xi \in \nabla_j^1} \left| (\tilde{\mathcal{M}} \tilde{\mathbf{a}})_{(j,\xi)} \right|^{p/r} \right]^{q/p} \right)^{q/r} \end{aligned}$$

which can be bounded from above by

$$\left\| \widetilde{M} \widetilde{\mathbf{a}} \Big| b_{p/r, q/r}^0(\nabla^1) \right\| \lesssim \left\| \widetilde{\mathbf{a}} \Big| b_{p/r, q/r}^0(\nabla^0) \right\| = \left\| \mathbf{a} \Big| b_{p,q}^0(\nabla^0) \right\|^r.$$

Here the last estimate follows from Step 1, since  $\widetilde{\mathcal{M}} \in \text{ad}_{p/r}^0$  and  $1 < p/r < \infty$ . Clearly, the same is true also for  $q = \infty$ . Thus, we have shown that  $\mathcal{M} \in \text{ad}_p^0$  implies the continuity of  $M: b_{p,q}^0(\nabla^0) \rightarrow b_{p,q}^0(\nabla^1)$ , as claimed.

*Step 3 (case  $\alpha_i \neq 0$ ).* Following [23] we note that the result for the case  $\alpha_i \neq 0$ ,  $i \in \{0, 1\}$ , can be reduced to the assertion for  $\alpha_0 = \alpha_1 = 0$  as follows: Obviously, for  $i \in \{0, 1\}$ , we have

$$\mathbf{a} = (a_{(k,\eta)})_{(k,\eta) \in \nabla^i} \in b_{p,q}^{\alpha_i}(\nabla^i) \quad \text{if and only if} \quad \widetilde{\mathbf{a}} := (2^{k\alpha_i} a_{(k,\eta)})_{(k,\eta) \in \nabla^i} \in b_{p,q}^0(\nabla^i)$$

with  $\left\| \mathbf{a} \Big| b_{p,q}^{\alpha_i}(\nabla^i) \right\| = \left\| \widetilde{\mathbf{a}} \Big| b_{p,q}^0(\nabla^i) \right\|$ . Moreover,  $\mathcal{M} = \{m_{(j,\xi),(k,\eta)}\}_{(j,\xi) \in \nabla^1, (k,\eta) \in \nabla^0} \in \text{ad}_p^{\alpha_0, \alpha_1}$  if and only if  $\widetilde{\mathcal{M}} := \{2^{j\alpha_1 - k\alpha_0} m_{(j,\xi),(k,\eta)}\}_{(j,\xi) \in \nabla^1, (k,\eta) \in \nabla^0}$  belongs to  $\text{ad}_p^0$ . Since  $\left\| M \mathbf{a} \Big| b_{p,q}^{\alpha_1}(\nabla^1) \right\|$  clearly equals  $\left\| \widetilde{M} \widetilde{\mathbf{a}} \Big| b_{p,q}^0(\nabla^1) \right\|$ , the linear operator  $M: b_{p,q}^{\alpha_0}(\nabla^0) \rightarrow b_{p,q}^{\alpha_1}(\nabla^1)$  with matrix  $\mathcal{M}$  is bounded if and only if  $\widetilde{M}: b_{p,q}^0(\nabla^0) \rightarrow b_{p,q}^0(\nabla^1)$  with matrix  $\widetilde{\mathcal{M}}$  is continuous. As this argument holds for every  $p$  and  $q$ , the proof is complete.  $\blacksquare$

### 3 Besov-type spaces based on wavelet expansions

In this section we turn to function spaces. In particular, here we are going to extend our notion of Besov-type spaces established in [13] to a fairly general setting. These (quasi-)Banach spaces are subsets of the space of all square-integrable functions defined on some domain or manifold  $\Gamma$ . In view of the applications we have in mind, we are especially interested in bounded manifolds which admit a decomposition into smooth parametric images of the unit cube in  $d$  spatial dimensions, since domains of this type are widely used in practice, e.g., in Computer Aided Geometric Design (CAGD). Moreover, as explained in the introduction, they are well-suited for the efficient numerical treatment of operator equations using FEM or BEM schemes based on multiscale analysis techniques. Consequently, in what follows we will focus on biorthogonal wavelet systems  $\Psi$  (as they were constructed and analyzed, e.g., in [2, 3, 18, 27]) for such patchwise smooth manifolds  $\Gamma$ . In the Besov-type spaces  $B_{\Psi,q}^\alpha(L_p(\Gamma))$  we then collect all those  $L_2(\Gamma)$ -functions, whose sequence of expansion coefficients w.r.t. some fixed basis  $\Psi$  decays sufficiently fast (i.e., belongs to the space  $b_{p,q}^\alpha(\nabla)$  introduced in Definition 2.3). Recently, it has been demonstrated that these Besov-type function spaces naturally arise in the analysis of adaptive numerical methods for operator equations on manifolds; the smoothness of the solutions (measured in these scales) determines

the rate of their best  $n$ -term (wavelet) approximation which in turn serves as a benchmark for the performance of ideal adaptive algorithms. For details we refer to [13], as well as to the introduction of the current paper, and to the references given therein.

In Section 3.1 below we describe the setting for the domains or manifolds under consideration in detail. Afterwards we recall some fundamental ideas from the field of multiscale analysis and review basic features of the three special wavelet constructions on manifolds we are going to deal with. Finally, Section 3.3 is concerned with the definition of Besov-type spaces based on wavelet expansions, as well as with some of their theoretical properties which are relevant for practical applications.

### 3.1 Domain decomposition and representation of geometry

When it comes to applications such as, e.g., the numerical treatment of integral equations defined on complicated geometries, often the following setting is assumed.

Given natural numbers  $m$  and  $d$  with  $d \leq m$ , let  $\Gamma$  denote a bounded  $d$ -dimensional manifold in  $\mathbb{R}^m$  with or without a boundary. We assume that  $\Gamma$  is at least globally Lipschitz continuous and admits a decomposition

$$\overline{\Gamma} = \bigcup_{i=1}^N \overline{\Gamma}_i \quad (9)$$

into finitely-many, essentially disjoint *patches*  $\Gamma_i$ , i.e.,  $\Gamma_i \cap \Gamma_j = \emptyset$  for all  $i \neq j$ . In addition, we assume that these patches are given as smooth parametric images of the  $d$ -dimensional unit cube which will serve as a reference domain. That is, we assume

$$\overline{\Gamma}_i = \kappa_i([0, 1]^d),$$

where for each  $i = 1, \dots, N$  the function  $\kappa_i: \mathbb{R}^d \rightarrow \mathbb{R}^m$  is supposed to be sufficiently regular. Moreover, the splitting of  $\Gamma$  needs to be conforming in the sense that for all  $i \neq j$  the pullback of the intersection  $\overline{\Gamma}_i \cap \overline{\Gamma}_j$  is either empty or a lower dimensional face of  $[0, 1]^d$ . In the latter case the set  $\overline{\Gamma}_i \cap \overline{\Gamma}_j$  is called *interface* between the patches  $\Gamma_i$  and  $\Gamma_j$ . Finally, we assume that the parametrizations  $\kappa_i$  are chosen in a way that for every interface there exists a permutation  $\pi_{i,j}$  such that

$$\kappa_j \circ \pi_{i,j} \circ \kappa_i^{-1} = \text{Id} \quad \text{on} \quad \overline{\Gamma}_i \cap \overline{\Gamma}_j,$$

where  $\text{Id}$  denotes the identical mapping and  $\pi_{i,j}(\mathbf{x}) := (x_{\pi_{i,j}(1)}, \dots, x_{\pi_{i,j}(d)})$  for  $\mathbf{x} = (x_1, \dots, x_d)$  in  $[0, 1]^d$ .

For the remainder of this paper a domain or manifold which meets all these requirements is said to be *decomposable* or *patchwise smooth*.

**Example 3.1.** Practically relevant examples for the manifolds under consideration are given by surfaces of bounded polyhedra in three spatial dimensions, i.e.,  $\Gamma = \partial\Omega$  with  $\Omega \subset \mathbb{R}^3$ . The reader may think of, e.g., Fichera's corner  $\Omega = [-1, 1]^3 \setminus [0, 1]^3$  which often serves as a model domain for numerical simulations. (Here the reentrant corner causes a singularity in the solutions to large classes of operator equations that is typical for problems on non-smooth domains.)  $\square$

**Remark 3.2.** We stress the fact that although our setting is tailored to handle boundary integral equations defined on (two-dimensional) closed surfaces, it covers open manifolds and bounded domains (of arbitrary dimension) as well. Thus, in principle the approach given here is suitable also for the treatment of boundary value problems involving partial differential operators.  $\square$

### 3.2 Multiresolution analysis and biorthogonal wavelets on patchwise smooth manifolds

One powerful tool to construct approximate solutions to operator equations defined on decomposable domains or manifolds in the sense the previous section is given by (adaptive) wavelet methods. In this approach the equation under consideration is discretized using a suitable set of basis functions stemming from a multiresolution analysis (a precise definition is given below). Then truncated versions of the resulting infinite linear system are solved. The attractive features of wavelets (such as smoothness, cancellation, and support properties which together imply the needed compression) combined with adaptive refinement and coarsening strategies finally yield an efficient algorithm. However, the construction of wavelet bases on patchwise smooth manifolds  $\Gamma$  is far away from being trivial as we shall now explain.

Let us assume for a moment that  $\Gamma$  denotes an arbitrary set (equipped with some metric) which additionally allows the definition of  $L_2(\Gamma)$ , the space of (equivalence classes of) square-integrable functions  $f: \Gamma \rightarrow \mathbb{C}$ , equipped with some inner product  $\langle \cdot, \cdot \rangle$ . Moreover, we assume to be given a *multiresolution analysis* (MRA) for this space, i.e., a sequence  $\mathcal{V} = (V_j)_{j \in \mathbb{N}}$  of closed linear subspaces of  $L_2(\Gamma)$  which satisfies

$$V_j \subset V_{j+1}, \quad j \in \mathbb{N}, \quad \text{and} \quad \overline{\bigcup_{j \in \mathbb{N}} V_j}^{L_2(\Gamma)} = L_2(\Gamma).$$

Now the main idea in multiscale analysis is to find a suitable *system of wavelet type*  $\Psi^\Gamma := \bigcup_{j \in \mathbb{N}_0} \Psi_j^\Gamma \subset L_2(\Gamma)$  such that the functions at *level*  $j$  span some complement of  $V_j$  in  $V_{j+1}$ .

Given any countable subset  $X \subset L_2(\Gamma)$  we let  $S(X)$  denote  $\overline{\text{span } X}^{L_2(\Gamma)}$ . Then we assume that

$$V_1 = S(\Psi_0^\Gamma) \quad \text{and} \quad V_{j+1} = V_j \oplus S(\Psi_j^\Gamma), \quad j \in \mathbb{N},$$

where, for every  $j \in \mathbb{N}_0$ ,  $\Psi_j^\Gamma := \{\psi_{j,\xi}^\Gamma \mid \xi \in \nabla_j^\Psi\}$  is indexed by some set  $\nabla_j^\Psi$  such that the sequence  $\nabla^\Psi := (\nabla_j^\Psi)_{j \in \mathbb{N}_0}$  forms a multiscale grid for  $\Gamma$  in the sense of Definition 2.1. Of course it would be favorable if  $\Psi^\Gamma$  would constitute an orthonormal basis for  $L_2(\Gamma)$ , but, in practice, such bases are not always feasible. However, as we shall see, there exist *biorthogonal* constructions which retain most of the desired properties of orthonormal bases.

In the (more flexible) biorthogonal setting a second multiresolution analysis  $\tilde{\mathcal{V}} = (\tilde{V}_j)_{j \in \mathbb{N}}$  of  $L_2(\Gamma)$ , together with a corresponding system of wavelet type  $\tilde{\Psi}^\Gamma$  (again indexed by  $\nabla^\Psi$ ), is needed such that the following duality w.r.t. the inner product  $\langle \cdot, \cdot \rangle$  holds for all  $j \in \mathbb{N}$ :

$$\tilde{V}_j \perp S(\Psi_j^\Gamma) \quad \text{and} \quad V_j \perp S(\tilde{\Psi}_j^\Gamma).$$

It then follows that both the systems  $\Psi^\Gamma$  and  $\tilde{\Psi}^\Gamma$  form (Schauder-) bases of  $L_2(\Gamma)$  and that they are biorthogonal in the sense that

$$\langle \psi_{j,\xi}^\Gamma, \tilde{\psi}_{k,\eta}^\Gamma \rangle = \begin{cases} 1, & \text{if } j = k \text{ and } \xi = \eta, \\ 0, & \text{otherwise.} \end{cases}$$

Finally, under suitable conditions, we can assume that  $\Psi^\Gamma$  and  $\tilde{\Psi}^\Gamma$  even form biorthogonal *Riesz bases* for  $L_2(\Gamma)$ , i.e., every  $u \in L_2(\Gamma)$  can be written as

$$u = \sum_{(j,\xi) \in \nabla^\Psi} \langle u, \tilde{\psi}_{j,\xi}^\Gamma \rangle \psi_{j,\xi}^\Gamma = \sum_{(j,\xi) \in \nabla^\Psi} \langle u, \psi_{j,\xi}^\Gamma \rangle \tilde{\psi}_{j,\xi}^\Gamma \quad (10)$$

and we have the norm equivalences

$$\|u\|_{L_2(\Gamma)} \sim \left\| \left( \langle u, \tilde{\psi}_{j,\xi}^\Gamma \rangle \right)_{(j,\xi) \in \nabla^\Psi} \right\|_{\ell_2(\nabla^\Psi)} \sim \left\| \left( \langle u, \psi_{j,\xi}^\Gamma \rangle \right)_{(j,\xi) \in \nabla^\Psi} \right\|_{\ell_2(\nabla^\Psi)}. \quad (11)$$

Consequently, then all (primal and dual) wavelets are normalized in the sense that

$$\|\psi_{j,\xi}^\Gamma\|_{L_2(\Gamma)} \sim \|\tilde{\psi}_{j,\xi}^\Gamma\|_{L_2(\Gamma)} \sim 1, \quad (j, \xi) \in \nabla^\Psi. \quad (12)$$

For the rest of this paper we shall use  $\Psi$  as a shortcut for (bi-)orthogonal wavelet Riesz bases  $(\Psi^\Gamma, \tilde{\Psi}^\Gamma)$  on the set  $\Gamma$  with the properties just mentioned.



**Remark 3.3.** When dealing with the (more restrictive) orthogonal setting we simply assume that the primal and the dual MRA, as well as the corresponding systems of wavelet type, coincide, i.e., then  $\tilde{\mathcal{V}} = \mathcal{V}$  and  $\tilde{\Psi}^\Gamma = \Psi^\Gamma$ .  $\square$

Although these concepts of multiscale representations can be employed in a quite general framework, for the ease of presentation and in view of the applications we have in mind, we are mainly interested in sets  $\Gamma$  which meet the requirements of Section 3.1 in what follows. In particular, we only consider manifolds of finite diameter and explicitly exclude the case of unbounded domains in  $\mathbb{R}^d$ . For this setting a suitable inner product for  $L_2(\Gamma)$ , which is equivalent to the canonical one, is given by

$$\langle f, g \rangle := \sum_{i=1}^N \langle f \circ \kappa_i, g \circ \kappa_i \rangle_{L_2([0,1]^d)}, \quad f, g \in L_2(\Gamma), \quad (13)$$

because it allows to shift the challenging problem of constructing wavelets from the (possibly complicated) manifold  $\Gamma$  to the unit cube  $[0, 1]^d$ . Thanks to the tensor product structure of this reference domain, multivariate wavelets then can be easily deduced from univariate ones which in turn are constructed with the help of some dual pair  $(\theta, \tilde{\theta})$  of refinable functions on the real line. When required by the final application even (homogeneous) boundary conditions can be incorporated at this point; see, e.g., [17, 30].

An important family of underlying dual pairs is based on B-splines, as they allow very efficient point evaluation and quadrature routines:

$$(\theta, \tilde{\theta}) := ({}_D\theta, {}_{D,\tilde{D}}\tilde{\theta}), \quad \text{where } D, \tilde{D} \in \mathbb{N} \quad \text{with } \tilde{D} \geq D \quad \text{and } D + \tilde{D} \text{ even.}$$

Therein  ${}_D\theta$  denotes the  $D$ th-order centered cardinal B-spline and  ${}_{D,\tilde{D}}\tilde{\theta}$  is some compactly supported, refinable function which is exact of order  $\tilde{D}$ . Moreover, it can be checked that the regularity of  ${}_D\theta$  equals  $D - 1/2$  and  ${}_{D,\tilde{D}}\tilde{\theta}$  can be chosen in a way such that its regularity increases proportional with  $\tilde{D}$ , i.e.,

$$\gamma := \sup\{s > 0 \mid {}_D\theta \in H^s(\mathbb{R})\} = D - \frac{1}{2} \quad \text{and} \quad \tilde{\gamma} := \sup\left\{s > 0 \mid {}_{D,\tilde{D}}\tilde{\theta} \in H^s(\mathbb{R})\right\} \sim \tilde{D}.$$

By now there exist several constructions that use the idea of lifting wavelets from the cube to the patches  $\Gamma_i$  of the manifold under consideration. In the sequel we particularly focus on the prominent case of *composite wavelet bases* which were initially established by Dahmen and Schneider in [18] and further developed by Harbrecht and Stevenson [27]. Another important set of wavelets is due to Canuto, Tabacco and Urban [2, 3]. These three

constructions (which will be labeled by  $\Psi_{\text{DS}}$ ,  $\Psi_{\text{HS}}$ , and  $\Psi_{\text{CTU}}$ , respectively) mainly differ in the treatment of wavelets supported in the vicinity of the interfaces. Without going into details, we mention that here some “gluing” or “matching” procedure is necessary in order to finally obtain wavelets which are continuous across the interfaces and, at the same time, yield the compression properties required by applications.

For these three systems it has been shown that for all choices of construction parameters the resulting wavelets provide norm equivalences

$$\|u\|_{H^s(\Gamma)} \sim \left( \sum_{j=0}^{\infty} 2^{2js} \sum_{\xi \in \nabla_j^\Psi} \left| \langle u, \tilde{\psi}_{j,\xi}^\Gamma \rangle \right|^2 \right)^{1/2} \quad (14)$$

for the scale of classical Sobolev spaces  $H^s(\Gamma)$  in some (very limited) range of smoothness parameters  $s$  which does not depend on  $D$  and  $\tilde{D}$ ; see, e.g., [18, Thm. 4.6.1]. Moreover, for  $s \in (-\tilde{\gamma}, D - 1/2)$  the same equivalences hold for generalized spaces  $H_s(\Gamma)$  based on  $\langle \cdot, \cdot \rangle$  which coincide with  $H^s(\Gamma)$  provided that  $|s|$  is sufficiently small. This shows that these (Hilbert) spaces are characterized by *all* bases  $\Psi = (\Psi^\Gamma, \tilde{\Psi}^\Gamma)$  under consideration as long as their construction parameters  $D^\Psi$  and  $\tilde{D}^\Psi$  are chosen sufficiently large. In Section 4 we are going to extend this assertion to a fairly large class of (quasi-)Banach spaces: We show that, in the sense of equivalent (quasi-)norms, any two wavelet systems  $\Psi, \Phi \in \{\Psi_{\text{DS}}, \Psi_{\text{HS}}, \Psi_{\text{CTU}}\}$  generate the same Besov-type spaces (see Definition 3.5 below)

$$B_{\Psi,q}^\alpha(L_p(\Gamma)) = B_{\Phi,q}^\alpha(L_p(\Gamma))$$

provided that the smoothness of the space  $\alpha$  is smaller than some quantity depending on  $D^\Psi$  and  $D^\Phi \in \mathbb{N}$ . To prove this we will have to bound the inner products  $\langle \psi_{k,\eta}^\Gamma, \tilde{\phi}_{j,\xi}^\Gamma \rangle$  subject to the relation of  $(j, \xi) \in \nabla^\Phi$  and  $(k, \eta) \in \nabla^\Psi$  to each other. For this purpose, the following properties shared by all the three bases of interest will be useful. As their proof is quite technical, we postpone it to the appendix; see Section 5.2.

**Lemma 3.4.** *For a decomposable  $d$ -dim. manifold  $\Gamma$  let  $\Psi = (\Psi^\Gamma, \tilde{\Psi}^\Gamma) \in \{\Psi_{\text{DS}}, \Psi_{\text{HS}}, \Psi_{\text{CTU}}\}$  denote a wavelet basis (as constructed in [18], [27], or [2, 3]) indexed by some multiscale grid  $\nabla^\Psi = (\nabla_j^\Psi)_{j \in \mathbb{N}_0}$  for  $\Gamma$ . Then for all  $j \in \mathbb{N}_0$  and each  $\xi = (y, t) \in \nabla_j^\Psi \subset \Gamma \times \mathcal{T}$  we have that*

$$(P1) \quad y \in \text{supp } \psi_{j,\xi}^\Gamma \cap \text{supp } \tilde{\psi}_{j,\xi}^\Gamma,$$

$$(P2) \quad \text{diam}(\text{supp } \psi_{j,\xi}^\Gamma) \sim \text{diam}(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma) \sim 2^{-j}, \text{ and}$$

(P3) there exist cubes  $\tilde{C}_{j,\xi}^i \subset [0, 1]^d$  with  $|\tilde{C}_{j,\xi}^i| \lesssim 2^{-jd}$  and  $\kappa_i^{-1}(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma \cap \Gamma_i) \subseteq \tilde{C}_{j,\xi}^i$  such that for every  $s \in (d/2, D^\Psi]$  and all functions  $f: \Gamma \rightarrow \mathbb{C}$  it holds

$$\left| \langle f, \tilde{\psi}_{j,\xi}^\Gamma \rangle \right| \lesssim \sum_{i=1}^N 2^{-js} |f \circ \kappa_i|_{H^s(\tilde{C}_{j,\xi}^i)}, \quad (15)$$

provided that the right-hand side is finite. Moreover, a completely analogous statement holds true for cubes  $C_{j,\xi}^i$  when  $\tilde{\psi}_{j,\xi}^\Gamma$  and  $D^\Psi$  are replaced by  $\psi_{j,\xi}^\Gamma$  and  $\tilde{D}^\Psi$ , respectively.

Since particularly the property (P3) will be essential in the following, let us add some comments on it: Roughly speaking, the estimate (15) says that the expansion coefficients of functions  $f$  on  $\Gamma$  which correspond to wavelets  $\psi_{j,\xi}^\Gamma$  supported on more than one patch  $\Gamma_i$  are bounded by the *patchwise* (Sobolev) regularity of the pullbacks of  $f$  to the unit cube. The lower bound  $d/2$  for  $s$  is due to the fact that the wavelet constructions under consideration incorporate terms which involve function evaluations at the interfaces, whereas the upper bound  $D^\Psi$  is implied by the degree of polynomial exactness of the underlying scaling functions.

### 3.3 Definition and properties of Besov-type function spaces

Besov spaces essentially generalize the concept of Sobolev spaces. On  $\mathbb{R}^d$  they are typically defined using harmonic analysis, finite differences, moduli of smoothness, or interpolation. Characteristics such as embeddings, interpolation results, or approximation properties of these scales then require deep proofs within the classical theory of function spaces. Often they are obtained by reducing the assertion of interest to the level of sequence spaces by means of characterizations in terms of building blocks (atoms, local means, quarks, or wavelets). To mention at least a few references the interested reader is referred to the monographs [31, 33], as well as to the articles [21, 23, 28].

As outlined in [13], the definition of Besov spaces on manifolds deserves some care: When following the usual approach based on local charts the smoothness of the spaces then is limited by the global regularity of the underlying manifold. On the other hand, the theoretical analysis of adaptive algorithms for problems defined on (patchwise smooth) manifolds naturally requires higher-order smoothness spaces of Besov type. Therefore, in [13, Def. 4.1] we introduced a notion of Besov-type spaces  $B_{\Psi,q}^\alpha(L_p(\Gamma))$  on specific two-dimensional closed manifolds such as boundaries of certain polyhedral domains  $\Omega \subset \mathbb{R}^3$ . The definition was based on expansions w.r.t. special wavelets bases  $\Psi$  that satisfy a number of additional properties; see [13] for details. Now we are going to extend this definition to a much more general setting: Besides enlarging the range of admissible parameters, here we significantly weaken

the assumptions on the underlying set  $\Gamma$ , as well as on the wavelet bases used in the construction of the spaces. As illustrated by Example 3.7 below, the subsequent definition thus covers a wide variety of function spaces on bounded or unbounded, smooth or non-smooth domains and manifolds of arbitrary dimension.

**Definition 3.5.** Let  $\Psi = (\Psi^\Gamma, \tilde{\Psi}^\Gamma)$  denote any (bi-)orthogonal wavelet Riesz basis for  $L_2(\Gamma)$  w.r.t. some inner product  $\langle \cdot, \cdot \rangle$  which is indexed by a multiscale grid  $\nabla^\Psi = (\nabla_j^\Psi)_{j \in \mathbb{N}_0}$  of dimension  $d \in \mathbb{N}$  for  $\Gamma$ . Then

(i) the tuple  $(\alpha, p, q)$  is said to be admissible if

$$0 < p \begin{cases} < \infty & \text{when } \nabla^\Psi \text{ satisfies (A4a),} \\ \leq 2 & \text{when } \nabla^\Psi \text{ satisfies (A4b),} \end{cases}$$

and if one of the following conditions applies:

- )  $\alpha > d \cdot \max\left\{0, \frac{1}{p} - \frac{1}{2}\right\}$  and  $0 < q \leq \infty$ ,
- )  $\alpha = d \cdot \max\left\{0, \frac{1}{p} - \frac{1}{2}\right\}$  and  $0 < q \leq 2$ .

(ii) for any admissible parameter tuple  $(\alpha, p, q)$  let  $B_{\Psi, q}^\alpha(L_p(\Gamma))$  denote the collection of all complex-valued functions  $u \in L_2(\Gamma)$  such that the (quasi-)norm defined by

$$\|u\|_{B_{\Psi, q}^\alpha(L_p(\Gamma))} := \left\| \left( \left\langle u, \tilde{\psi}_{j, \xi}^\Gamma \right\rangle \right)_{(j, \xi) \in \nabla^\Psi} \right\|_{b_{p, q}^\alpha(\nabla^\Psi)}$$

is finite. Therein the sequence space  $b_{p, q}^\alpha(\nabla^\Psi)$  is defined as in Definition 2.3.

Some comments are in order.

**Remark 3.6.** Observe that due to our assumptions every  $u \in L_2(\Gamma)$  admits a unique expansion w.r.t. the primal wavelet system  $\Psi^\Gamma$ , where the corresponding sequence of coefficients belongs to  $\ell_2(\nabla^\Psi) = b_{2, 2}^0(\nabla^\Psi)$ ; cf. (10) and (11). On the other hand, Proposition 2.5 implies that  $b_{p, q}^\alpha(\nabla^\Psi) \hookrightarrow \ell_2(\nabla^\Psi)$  for all admissible parameter tuples. Therefore every function with finite  $B_{\Psi, q}^\alpha(L_p(\Gamma))$ -quasi-norm belongs to  $L_2(\Gamma)$ . In fact, we have  $B_{\Psi, q}^\alpha(L_p(\Gamma)) \hookrightarrow L_2(\Gamma)$ ; also compare Corollary 3.8 below.  $\square$

**Example 3.7.** Let us illustrate the flexibility of Definition 3.5 by means of the some examples:

- (i) Most importantly, Definition 3.5 covers spaces on  $d$ -dimensional manifolds which are patchwise smooth in the sense of Section 3.1. As exposed in Section 3.2, in this setting biorthogonality of functions on  $\Gamma$  usually is realized w.r.t. the patchwise inner product (13). Suitable wavelet systems are given by  $\Psi = (\Psi^\Gamma, \tilde{\Psi}^\Gamma) \in \{\Psi_{\text{DS}}, \Psi_{\text{HS}}, \Psi_{\text{CTU}}\}$ , as constructed in [18], [27], and [2, 3], respectively. Note that then no restriction on  $p$  is imposed as decomposable domains (9) are assumed to be bounded; cf. (A4a).
- (ii) Assume that  $\Gamma$  denotes a compact  $d$ -dimensional  $\mathcal{C}^\infty$  manifold. Then Besov spaces  $B_{p,q}^\alpha(\Gamma)$  of arbitrary smoothness are well-defined by lifting spaces of distributions on  $\mathbb{R}^d$  to  $\Gamma$  using local charts together with an (overlapping) resolution of unity; see, e.g., [34, Def. 5.1]. Without going into details, we state that (for the range of admissible parameter tuples) these spaces coincide with our spaces  $B_{\Psi,q}^\alpha(L_p(\Gamma))$  introduced in Definition 3.5, provided that the wavelet system under consideration satisfies additional requirements; cf. [34, Prop. 5.32]. For an elaborate discussion we refer to [34, Ch. 5].
- (iii) Finally, note that also classical Besov function spaces  $B_{p,q}^\alpha(\mathbb{R}^d)$  are covered. For this purpose, we may take a system of Daubechies wavelets which forms an orthogonal basis w.r.t. the canonical inner product on  $L_2(\mathbb{R}^d)$ . The coincidence of our Definition 3.5 with the usual definition based on Fourier techniques then is shown, e.g., in [33, Thm. 1.64]; see also the proof of Proposition 2.5 in Section 5.2.  $\square$

In the remainder of this section we briefly describe a couple of properties satisfied by the scale of function spaces just introduced which yield attractive implications for practical applications, e.g., in the context of regularity studies of operator equations. For details we again refer to [13]. To begin with, we note that (formally) the spaces constructed in Definition 3.5 depend on the concrete choice of the wavelet basis  $\Psi$ . As already mentioned, in Section 4 below we will show that under quite natural conditions all wavelet systems under consideration actually lead to the same Besov-type spaces.

From the properties of the sequence spaces  $b_{p,q}^\alpha(\nabla^\Psi)$  it immediately follows that all spaces  $B_{\Psi,q}^\alpha(L_p(\Gamma))$  are quasi-Banach spaces; cf. Remark 2.4. Moreover, they are Banach spaces if and only if  $\min\{p, q\} \geq 1$  and Hilbert spaces if and only if  $p = q = 2$ . In fact, for small smoothness parameters  $\alpha = s \in [0, \min\{3/2, s_\Gamma\})$  and  $p = q = 2$  our Besov-type spaces coincide with the classical Sobolev Hilbert spaces  $H^s(\Gamma)$  (in the sense of equivalent norms), provided that a suitable wavelet basis is used; see, e.g., [18, Thm. 4.6.1], or [2, Cor. 5.7]. Here the number  $s_\Gamma$  is related to the smoothness of the underlying manifold  $\Gamma$ . The coincidence then simply follows from the fact that the right-hand side of (14) equals  $\left\| (\langle u, \tilde{\psi}_{j,\xi}^\Gamma \rangle)_{(j,\xi) \in \nabla^\Psi} \right\|_{b_{2,2}^s(\nabla^\Psi)}$  which in turn defines the norm of  $u$  in  $B_{\Psi,2}^s(L_2(\Gamma))$ .

Furthermore, Proposition 2.5 (together with Remark 3.6) implies the subsequent characterization of embeddings between Besov-type spaces which is listed here for the sake of

completeness.

**Corollary 3.8** (Standard embeddings). *Choose  $\Psi$ ,  $\nabla^\Psi$ , and  $\Gamma$  as in Definition 3.5. Moreover, for  $\alpha, \gamma \in \mathbb{R}$ , let  $(\alpha + \gamma, p_0, q_0)$  and  $(\alpha, p_1, q_1)$  denote admissible parameter tuples. If  $\Gamma$  is bounded (i.e.,  $\nabla^\Psi$  satisfies (A4a)) then we have the continuous embedding*

$$B_{\Psi, q_0}^{\alpha+\gamma}(L_{p_0}(\Gamma)) \hookrightarrow B_{\Psi, q_1}^\alpha(L_{p_1}(\Gamma))$$

if and only if one of the following conditions applies:

- )  $\gamma > d \cdot \max\left\{0, \frac{1}{p_0} - \frac{1}{p_1}\right\},$
- )  $\gamma = d \cdot \max\left\{0, \frac{1}{p_0} - \frac{1}{p_1}\right\} \quad \text{and} \quad q_0 \leq q_1.$

Furthermore, if  $\Gamma$  is unbounded (i.e.,  $\nabla^\Psi$  satisfies (A4b) instead of (A4a)) then a corresponding characterization holds true with the additional condition  $p_0 \leq p_1$ .

The embeddings stated in Corollary 3.8 can be illustrated by *DeVore-Triebel diagrams*; see Figure 1. Therein the solid lines, starting from the point  $(1/2; 0)$  which corresponds to the space  $L_2(\Gamma) = B_{\Psi, 2}^0(L_2(\Gamma))$ , describe the boundaries of the respective areas of admissible parameters; cf. Definition 3.5(i). In both cases these areas are limited at the right-hand side by the so-called *adaptivity scale of Besov-type spaces*  $B_{\Psi, \tau}^{\alpha_\tau}(L_\tau(\Gamma))$ , where  $\tau$  and  $\alpha_\tau$  are related via (3). Moreover, the shaded regions refer to all spaces which are embedded into  $H^s(\Gamma) = B_{\Psi, 2}^s(L_2(\Gamma))$ , whereas the arrows indicate limiting cases for possible embeddings of the space  $B_{\Psi, q_0}^{\alpha+\gamma}(L_{p_0}(\Gamma))$ . Restrictions imposed by the fine indices  $q$  are not visualized.

When it comes to applications on bounded domains or manifolds, approximation properties such as best  $n$ -term rates are of particular interest. Without going into details, let us recall that roughly speaking the numbers  $\sigma_n(F; \mathcal{B}, G)$ ,  $n \in \mathbb{N}_0$ , describe the minimal error of approximating the embedding  $F \hookrightarrow G$  by means of finite linear combinations of elements from the dictionary  $\mathcal{B}$ . For an exact definition we refer to [13, Sect. 4.2]. There also a proof (based on results shown in [11, 25]) of the next proposition for  $d = 2$  can be found. The arguments easily carry over to the general case discussed here.

**Proposition 3.9** (Best  $n$ -term approximation on bounded manifolds). *Choose  $\Psi$ ,  $\nabla^\Psi$ , and  $\Gamma$  as in Definition 3.5 and assume  $\Gamma$  to be bounded, i.e., suppose that  $\nabla^\Psi$  satisfies (A4a). Moreover, for  $\alpha, \gamma \in \mathbb{R}$ , let  $(\alpha + \gamma, p_0, q_0)$  and  $(\alpha, p_1, q_1)$  denote admissible parameter tuples. Then*

- )  $\gamma > d \cdot \max\left\{0, \frac{1}{p_0} - \frac{1}{p_1}\right\}$  implies

$$\sigma_n(B_{\Psi, q_0}^{\alpha+\gamma}(L_{p_0}(\Gamma)); \Psi^\Gamma, B_{\Psi, q_1}^\alpha(L_{p_1}(\Gamma))) \sim n^{-\gamma/d},$$

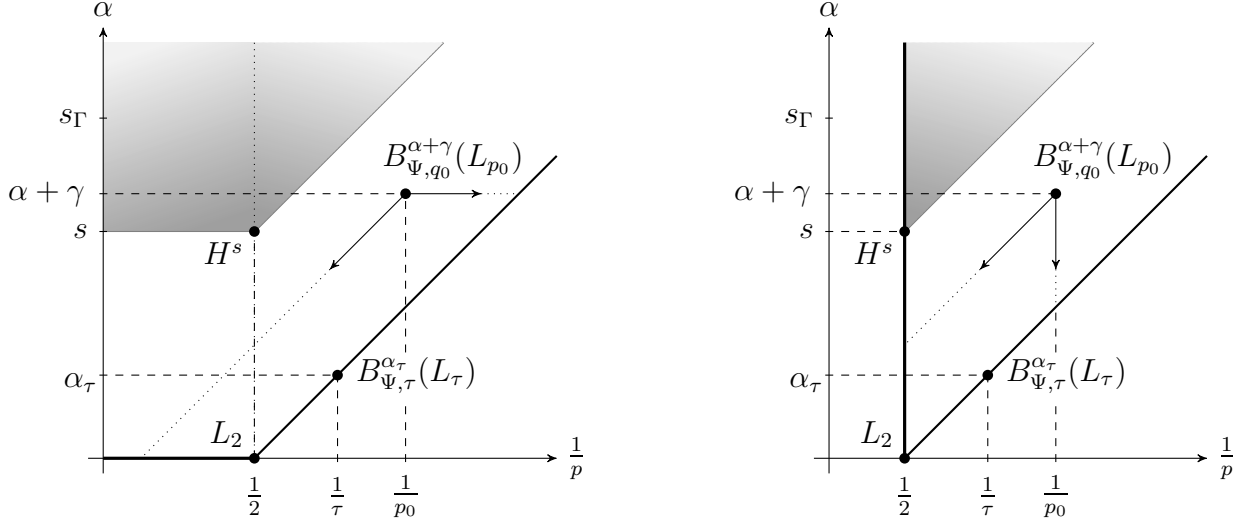


Figure 1: Embeddings for Besov-type spaces  $B_{\Psi,q}^\alpha(L_p(\Gamma))$  on bounded (left) and on unbounded (right) domains or manifolds  $\Gamma$ , resp.; cf. Corollary 3.8.

- )  $\gamma = d \cdot \max\left\{0, \frac{1}{p_0} - \frac{1}{p_1}\right\}$  and  $q_0 \leq q_1$  implies

$$\sigma_n(B_{\Psi,q_0}^{\alpha+\gamma}(L_{p_0}(\Gamma)); \Psi^\Gamma, B_{\Psi,q_1}^\alpha(L_{p_1}(\Gamma))) \sim n^{-\min\{\gamma/d, 1/q_0 - 1/q_1\}}.$$

**Remark 3.10.** Note that (using results stated in [25]) a corresponding characterization (but with quite different rates of convergence!) can be derived easily also for spaces on unbounded sets  $\Gamma$ . As for applications bounded domains or manifolds, respectively, are much more important, we will not follow this line of research here. Indeed, as exposed already in the introduction, the quantity  $\sigma_n(B_{\Psi,\tau}^{\alpha_\tau}(L_\tau(\Gamma)); \Psi^\Gamma, B_{\Psi,2}^0(L_2(\Gamma)))$  with  $(\alpha_\tau, \tau)$  as in (3) serves as a benchmark for the performance of (ideal) adaptive algorithms that use at most  $n$  wavelets from the dictionary  $\Psi^\Gamma$  and provide an approximation in the norm of  $L_2(\Gamma) = B_{\Psi,2}^0(L_2(\Gamma))$ . The reason is that, on the one hand, due to Proposition 3.9, the best  $n$ -term approximation rates linearly depend on the difference in smoothness and, on the other hand, the spaces  $B_{\Psi,\tau}^{\alpha_\tau}(L_\tau(\Gamma))$  from the adaptivity scale (3) provide the weakest norms among all (Besov-type) spaces of fixed regularity which are contained in  $L_2(\Gamma)$ ; cf. Corollary 3.8.  $\square$

Finally, besides many other interesting properties which are typical for classical Besov spaces (e.g., defined via harmonic analysis), our scale of Besov-type spaces  $B_{\Psi,q}^\alpha(L_p(\Gamma))$  satisfies the following interpolation assertions w.r.t the *real* and the (extended) *complex method* which we denote by  $(\cdot, \cdot)_{\Theta,q}$  and  $[\cdot, \cdot]_\Theta$ , respectively. For a comprehensive treatment of interpolation of (quasi-)Banach spaces we refer to [1, 28, 35] and to the references therein.

**Proposition 3.11** (Interpolation). *Choose  $\Psi$ ,  $\nabla^\Psi$ , and  $\Gamma$  as in Definition 3.5 and let  $(\alpha_0, p_0, q_0)$  and  $(\alpha_1, p_1, q_1)$  denote admissible parameter tuples. For  $0 < \Theta < 1$  we set*

$$s_\Theta := (1 - \Theta) \alpha_0 + \Theta \alpha_1, \quad \frac{1}{p_\Theta} := \frac{1 - \Theta}{p_0} + \frac{\Theta}{p_1}, \quad \text{and} \quad \frac{1}{q_\Theta} := \frac{1 - \Theta}{q_0} + \frac{\Theta}{q_1}.$$

- ) *If  $\alpha_0 \neq \alpha_1$  and  $p = p_0 = p_1$  then for all  $0 < q \leq \infty$  and every  $0 < \Theta < 1$  we have*

$$\left( B_{\Psi, q_0}^{\alpha_0}(L_p(\Gamma)), B_{\Psi, q_1}^{\alpha_1}(L_p(\Gamma)) \right)_{\Theta, q} = B_{\Psi, q}^{s_\Theta}(L_p(\Gamma)).$$

- ) *If  $\min\{q_0, q_1\} < \infty$  then for all  $0 < \Theta < 1$  it holds*

$$\left[ B_{\Psi, q_0}^{\alpha_0}(L_{p_0}(\Gamma)), B_{\Psi, q_1}^{\alpha_1}(L_{p_1}(\Gamma)) \right]_\Theta = B_{\Psi, q_\Theta}^{s_\Theta}(L_{p_\Theta}(\Gamma)).$$

*Proof.* As it has been shown in [13, Prop. 4.5] (for the special case  $d = 2$ ), interpolation results for Besov-type spaces  $B_{\Psi, q}^\alpha(L_p(\Gamma))$  can be reduced to corresponding assertions for sequence spaces which in turn follow from interpolation properties of (classical) Besov spaces  $B_{p, q}^\alpha(\mathbb{R}^d)$  defined on the whole of  $\mathbb{R}^d$ . This type of arguments does not depend on the dimension and can be applied for all methods that fulfill the so-called interpolation property; cf. [13, Rem. 6.3]. Thus, in our case it suffices to refer to [35, Thm. 2.4.2(i)] for the real method and to [28, Thm. 9.1] for the (extended) complex method, respectively. ■

## 4 Change of basis embeddings for Besov-type spaces

As outlined above, our Definition 3.5 of Besov-type spaces  $B_{\Psi, q}^\alpha(L_p(\Gamma))$  formally depends on the concrete choice of the wavelet basis  $\Psi$  and its construction parameters  $D^\Psi$  and  $\tilde{D}^\Psi$ , respectively. In order to find conditions which imply that different bases  $\Psi$  and  $\Phi$  generate the same Besov-type space  $B_{\Psi, q}^\alpha(L_p(\Gamma)) = B_{\Phi, q}^\alpha(L_p(\Gamma))$  in the sense of equivalent (quasi-)norms, we now employ the theory of almost diagonal matrices developed in Section 2 to investigate properties which yield corresponding one-sided *change of basis embeddings*.

Note that, in general, different constructions of wavelet bases might accomplish (bi-)orthogonality w.r.t. different inner products. Indeed, depending on the desired properties we like to assemble, on patchwise smooth manifolds  $\Gamma$ , say, it is reasonable to construct wavelets which are biorthogonal with respect to  $\langle \cdot, \cdot \rangle$  as defined in (13), or w.r.t. the canonical scalar product  $\langle\langle \cdot, \cdot \rangle\rangle$  on  $L_2(\Gamma)$ . The following proposition addresses this issue, as it is stated in a quite general form.



**Proposition 4.1.** *Let  $\Psi$  and  $\Phi$  denote two wavelet Riesz bases for  $L_2(\Gamma)$  which are (bi-) orthogonal w.r.t. the inner products  $\langle \cdot, \cdot \rangle$  and  $\langle\langle \cdot, \cdot \rangle\rangle$ , respectively. Moreover, suppose that these bases are indexed by multiscale grids  $\nabla^\Psi$  and  $\nabla^\Phi$  for  $\Gamma$ , respectively, and assume that for some admissible parameter tuple  $(\alpha, p, q)$  the associated Gramian matrix satisfies*

$$\mathcal{M}_{\Psi \rightarrow \Phi} := \{m_{(j,\xi),(k,\eta)}\}_{(j,\xi) \in \nabla^\Phi, (k,\eta) \in \nabla^\Psi} = \left\{ \left\langle\left\langle \psi_{k,\eta}^\Gamma, \tilde{\phi}_{j,\xi}^\Gamma \right\rangle\right\rangle \right\}_{(j,\xi) \in \nabla^\Phi, (k,\eta) \in \nabla^\Psi} \in \text{ad}_p^\alpha. \quad (16)$$

Then  $B_{\Psi,q}^\alpha(L_p(\Gamma)) \hookrightarrow B_{\Phi,q}^\alpha(L_p(\Gamma))$ .

*Proof.* We essentially follow the lines of the proof of [23, Thm. 3.7]. By definition, every  $u \in B_{\Psi,q}^\alpha(L_p(\Gamma))$  can be expanded into

$$u = \sum_{k \in \mathbb{N}_0} \sum_{\eta \in \nabla_k^\Psi} a_{(k,\eta)} \psi_{k,\eta}^\Gamma \quad \text{with} \quad \mathbf{a} := (a_{(k,\eta)})_{(k,\eta) \in \nabla^\Psi} = \left( \left\langle u, \tilde{\psi}_{k,\eta}^\Gamma \right\rangle \right)_{(k,\eta) \in \nabla^\Psi} \in b_{p,q}^\alpha(\nabla^\Psi).$$

Note that, since  $B_{\Psi,q}^\alpha(L_p(\Gamma)) \hookrightarrow L_2(\Gamma)$  and  $\Phi = (\Phi^\Gamma, \tilde{\Phi}^\Gamma)$  is a  $\langle\langle \cdot, \cdot \rangle\rangle$ -biorthogonal Riesz basis for  $L_2(\Gamma)$ , the sequence  $\tilde{\mathbf{a}} := (\tilde{a}_{(j,\xi)})_{(j,\xi) \in \nabla^\Phi} := \left( \left\langle\left\langle u, \tilde{\phi}_{j,\xi}^\Gamma \right\rangle\right\rangle \right)_{(j,\xi) \in \nabla^\Phi}$  is well-defined. Moreover, it holds  $\tilde{\mathbf{a}} = M_{\Psi \rightarrow \Phi} \mathbf{a}$ . That is, for all  $j \in \mathbb{N}_0$  and  $\xi \in \nabla_j^\Phi$ , we have

$$\tilde{a}_{(j,\xi)} = \left\langle\left\langle \sum_{k \in \mathbb{N}_0} \sum_{\eta \in \nabla_k^\Psi} a_{(k,\eta)} \psi_{k,\eta}^\Gamma, \tilde{\phi}_{j,\xi}^\Gamma \right\rangle\right\rangle = \sum_{(k,\eta) \in \nabla^\Psi} m_{(j,\xi),(k,\eta)} a_{(k,\eta)} = (\mathcal{M}_{\Psi \rightarrow \Phi} \mathbf{a})_{(j,\xi)}.$$

Thus, from Theorem 2.9 it follows

$$\begin{aligned} \|u\|_{B_{\Phi,q}^\alpha(L_p(\Gamma))} &= \|\tilde{\mathbf{a}}\|_{b_{p,q}^\alpha(\nabla^\Phi)} \\ &= \|M_{\Psi \rightarrow \Phi} \mathbf{a}\|_{b_{p,q}^\alpha(\nabla^\Phi)} \\ &\lesssim \|\mathbf{a}\|_{b_{p,q}^\alpha(\nabla^\Psi)} \\ &= \|u\|_{B_{\Psi,q}^\alpha(L_p(\Gamma))}. \end{aligned}$$

This shows that the identity (mapping  $B_{\Psi,q}^\alpha(L_p(\Gamma))$  into  $B_{\Phi,q}^\alpha(L_p(\Gamma))$ ), induced by the operator  $M_{\Psi \rightarrow \Phi}: b_{p,q}^\alpha(\nabla^\Psi) \rightarrow b_{p,q}^\alpha(\nabla^\Phi)$  which in turn is represented by the matrix  $\mathcal{M}_{\Psi \rightarrow \Phi}$  defined in (16), indeed is continuous, as claimed.  $\blacksquare$

Next let us apply the general concept presented in Proposition 4.1 above to the practically relevant case of Besov-type spaces generated by wavelet bases  $\Psi, \Phi \in \{\Psi_{\text{DS}}, \Psi_{\text{HS}}, \Psi_{\text{CTU}}\}$  on patchwise smooth ( $d$ -dimensional) manifolds  $\Gamma$  in the sense of Section 3.1. As described in Section 3.2, all of these constructions are biorthogonal with respect to the same inner

product (13) and all of them are built up from univariate centered cardinal B-splines of order  $D^\Psi$  and  $D^\Phi$  with regularity  $\gamma^\Psi$  and  $\gamma^\Phi$ , respectively. Again the corresponding dual quantities are denoted by  $\tilde{D}^\Psi$ , etc. Then the combination of Proposition 4.1 with Lemma 3.4 implies the following result.

**Theorem 4.2.** *For a patchwise smooth  $d$ -dim. manifold  $\Gamma$  let  $\Psi, \Phi \in \{\Psi_{\text{DS}}, \Psi_{\text{HS}}, \Psi_{\text{CTU}}\}$  denote two wavelet bases as constructed in [18], [27], or [2, 3], respectively. Moreover, assume that their construction parameters satisfy*

$$\min\{D^\Phi, \tilde{\gamma}^\Phi, \tilde{D}^\Psi, \gamma^\Psi\} > d/2.$$

*Then for all admissible tuples of parameters  $(\alpha, p, q)$  with*

$$0 \leq \alpha < \min\{D^\Phi, \gamma^\Psi\} \quad (17)$$

*the continuous embedding  $B_{\Psi, q}^\alpha(L_p(\Gamma)) \hookrightarrow B_{\Phi, q}^\alpha(L_p(\Gamma))$  holds true.*

*Proof. Step 1.* In order to prove the claim we like to apply Proposition 4.1. Thus we have to show that the Gramian matrix (w.r.t the change of basis from  $\Psi$  to  $\Phi$ )

$$\mathcal{M}_{\Psi \rightarrow \Phi} := \left\{ \left\langle \psi_{k, \eta}^\Gamma, \tilde{\phi}_{j, \xi}^\Gamma \right\rangle \right\}_{(j, \xi) \in \nabla^\Phi, (k, \eta) \in \nabla^\Psi}$$

belongs to the class  $\text{ad}_p^\alpha$  (cf. Definition 2.7) for  $\alpha$  and  $p$  under consideration. Here  $\nabla^\Psi$  and  $\nabla^\Phi$  denote the associated multiscale grids of dimension  $d$  for  $\Gamma$  and  $\langle \cdot, \cdot \rangle$  is defined in (13). Due to the monotonicity of the classes  $\text{ad}_p^\alpha$  (see Remark 2.8(ii)) it suffices to consider the limiting case  $p = \tau = \tau(\alpha)$  with

$$\tau^{-1} := \frac{\alpha}{d} + \frac{1}{2} \quad (18)$$

and  $\alpha$  that satisfies (17). Furthermore, as we will show in Step 2 below, it follows from the support conditions (P1) and (P2) in Lemma 3.4 that  $[1 + \min\{2^k, 2^j\} \text{dist}(\xi, \eta)] \sim 1$ , so that it is enough to show that there exists  $\varepsilon > 0$  such that

$$\left| \left\langle \psi_{k, \eta}^\Gamma, \tilde{\phi}_{j, \xi}^\Gamma \right\rangle \right| \lesssim \min\{2^{-(j-k)(d/2+\alpha+\varepsilon)}, 2^{(j-k)(d/2-\alpha+\varepsilon+\sigma_\tau)}\} \quad \text{for all } (j, \xi) \in \nabla^\Phi, (k, \eta) \in \nabla^\Psi. \quad (19)$$

Afterwards, in Step 3, we complete the proof by showing that (19) is implied by (P3) in Lemma 3.4.

*Step 2.* For  $\zeta := (y, t) \in \Gamma \times \mathcal{T}$  and  $r > 0$  let  $B(\zeta, r) := \{y' \in \Gamma \mid \varrho_\Gamma(y', y) < r\}$  denote the open ball of radius  $r$  around  $y$  in  $\Gamma$ . Then (P1) and (P2) imply that

$$\text{supp } \psi_{k, \eta}^\Gamma \cap \text{supp } \tilde{\phi}_{j, \xi}^\Gamma \subseteq B(\eta, c' 2^{-k}) \cap B(\xi, c' 2^{-j})$$

for some  $c' > 0$ , all  $j, k \in \mathbb{N}_0$ , and every  $\xi \in \nabla_j^\Phi$ ,  $\eta \in \nabla_k^\Psi$ , respectively. Note that the latter intersection is empty if  $\text{dist}(\xi, \eta) > c'(2^{-k} + 2^{-j})$  which in turn shows that  $\langle \psi_{k,\eta}^\Gamma, \tilde{\phi}_{j,\xi}^\Gamma \rangle \neq 0$  only if

$$1 \leq 1 + \min\{2^k, 2^j\} \text{dist}(\xi, \eta) \leq c''$$

for some  $c'' \geq 1$  which does not depend on  $j$  and  $k$ . Therefore (4), i.e., membership of  $\mathcal{M}_{\Psi \rightarrow \Phi}$  in  $\text{ad}_p^\alpha$ , is equivalent to (19), as promised.

*Step 3.* We show (19). For this purpose, we note that

$$\sigma_\tau = d \cdot \max\left\{\frac{1}{\tau} - 1, 0\right\} = \begin{cases} 0, & \text{if } 1 \leq \tau \leq 2, \\ d/\tau - d, & \text{if } 0 < \tau < 1 \end{cases} = \begin{cases} 0, & \text{if } 0 \leq \alpha \leq d/2, \\ \alpha - d/2, & \text{if } d/2 < \alpha, \end{cases}$$

due to (18). This leads to the observation that

$$\frac{d}{2} - \alpha + \varepsilon + \sigma_\tau = \begin{cases} d/2 - \alpha + \varepsilon, & \text{if } 0 \leq \alpha \leq d/2 \\ \varepsilon, & \text{if } d/2 < \alpha \end{cases} \geq \varepsilon > 0$$

such that the proof of (19) naturally splits into the cases  $j \geq k$  and  $j < k$ . For  $j, k \in \mathbb{N}_0$  with  $j \geq k$  we apply the first part of (P3) in Lemma 3.4 for the basis  $\Phi$  and  $f := \psi_{k,\eta}^\Gamma$  with  $\eta \in \nabla_k^\Psi$ . Observe that the patchwise regularity of this primal wavelet is as large as the smoothness of the underlying univariate spline used for its construction. Hence, given  $i \in \{1, \dots, N\}$ , we conclude

$$|\psi_{k,\eta}^\Gamma \circ \kappa_i|_{H^s(\tilde{C}_{j,\xi}^i)} \lesssim 2^{ks} \left\| \psi_{k,\eta}^\Gamma \circ \kappa_i \right\|_{L_2(\tilde{C}_{j,\xi}^i)} \quad \text{for all } 0 \leq s < \gamma^\Psi.$$

For  $s \in \mathbb{N}$  this simply follows from the multiscale structure of the wavelets. The case  $s \notin \mathbb{N}$  can be derived using standard interpolation arguments. Furthermore,

$$\left\| \psi_{k,\eta}^\Gamma \circ \kappa_i \right\|_{L_2(\tilde{C}_{j,\xi}^i)}^2 = \int_{\tilde{C}_{j,\xi}^i} |\psi_{k,\eta}^\Gamma(\kappa_i(x))|^2 dx \leq \|\psi_{k,\eta}^\Gamma\|_{L_\infty(\Gamma)}^2 \cdot |\tilde{C}_{j,\xi}^i| \lesssim 2^{kd} \cdot 2^{-jd},$$

since  $\psi_{k,\eta}^\Gamma$  is  $L_2(\Gamma)$ -normalized; see (12). Combining the two last estimates with (15) thus gives

$$\left| \langle \psi_{k,\eta}^\Gamma, \tilde{\phi}_{j,\xi}^\Gamma \rangle \right| \lesssim \sum_{i=1}^N 2^{-js} 2^{ks} 2^{(k-j)d/2} \sim 2^{-(j-k)(d/2+s)}$$

for all  $d/2 < s \leq D^\Phi$  with  $s < \gamma^\Psi$  and  $\xi \in \nabla_j^\Phi$ ,  $\eta \in \nabla_k^\Psi$  with  $j \geq k$  in  $\mathbb{N}_0$ . Note that, due to the assumption (17), we can find some  $s$  in this range which is strictly larger than  $\alpha$ .

Choosing  $\varepsilon > 0$  sufficiently small then yields  $2^{-(j-k)(d/2+s)} \leq 2^{-(j-k)(d/2+\alpha+\varepsilon)}$  which finally shows (19) for  $j \geq k$ .

We are left with the case  $j < k$ . Using the second part of (P3) in Lemma 3.4 (for the basis  $\Psi$ , the index  $\eta \in \nabla_k^\Psi$ , and  $f := \tilde{\phi}_{j,\xi}^\Gamma$  with  $\xi \in \nabla_j^\Phi$ ) together with the same arguments as before, we deduce the bound

$$\left| \left\langle \psi_{k,\eta}^\Gamma, \tilde{\phi}_{j,\xi}^\Gamma \right\rangle \right| = \left| \left\langle \tilde{\phi}_{j,\xi}^\Gamma, \psi_{k,\eta}^\Gamma \right\rangle \right| \lesssim 2^{-(k-j)(d/2+s)}$$

for all  $d/2 < s \leq \tilde{D}^\Psi$  with  $s < \tilde{\gamma}^\Phi$  and  $j < k$  in  $\mathbb{N}_0$ . Observe that for every such  $s$  we have the estimate  $2^{-(k-j)(d/2+s)} \leq 2^{(j-k)(d/2-\alpha+\varepsilon+\sigma_\tau)}$ , provided that  $\varepsilon > 0$  is chosen small enough. Therefore (19) also holds for  $j < k$  which completes the proof.  $\blacksquare$

As an immediate consequence of Theorem 4.2 we conclude the main result of this paper. It states that all wavelet bases  $\Psi, \Phi \in \{\Psi_{\text{DS}}, \Psi_{\text{HS}}, \Psi_{\text{CTU}}\}$  induce the same Besov-type spaces  $B_{\Psi,q}^\alpha(L_p(\Gamma)) = B_{\Phi,q}^\alpha(L_p(\Gamma))$  on patchwise smooth manifolds  $\Gamma$  (in the sense of equivalent quasi-norms), provided that the primal sets of wavelets are of sufficiently large order of cancellation and regularity compared to the smoothness parameter  $\alpha$  of the space.

**Theorem 4.3.** *Given some  $d$ -dimensional manifold  $\Gamma$  which is patchwise smooth in the sense of Section 3.1 let  $\Psi = (\Psi^\Gamma, \tilde{\Psi}^\Gamma)$  and  $\Phi = (\Phi^\Gamma, \tilde{\Phi}^\Gamma)$  denote two wavelet bases from  $\{\Psi_{\text{DS}}, \Psi_{\text{HS}}, \Psi_{\text{CTU}}\}$  as constructed in [18], [27], or [2, 3], respectively, and assume that their construction parameters satisfy*

$$\min\left\{D^\Psi, \tilde{D}^\Psi, \gamma^\Psi, \tilde{\gamma}^\Psi, D^\Phi, \tilde{D}^\Phi, \gamma^\Phi, \tilde{\gamma}^\Phi\right\} > d/2.$$

*Then, for all admissible tuples of parameters  $(\alpha, p, q)$  with*

$$0 \leq \alpha < \min\{D^\Psi, D^\Phi, \gamma^\Psi, \gamma^\Phi\},$$

*it holds  $B_{\Psi,q}^\alpha(L_p(\Gamma)) = B_{\Phi,q}^\alpha(L_p(\Gamma))$  in the sense of equivalent (quasi-)norms.*

## 5 Appendix

### 5.1 Auxiliary assertions

This part of the appendix is concerned with auxiliary (technical) assertions that are needed in our proofs. We start with a result which can be easily derived from [32, Lem. 2].

**Lemma 5.1.** *Let  $0 < r < \infty$ , as well as  $0 < q \leq \infty$ , and  $\mathbf{x} := (x_k)_{k \in \mathbb{N}_0} \in \ell_q(\mathbb{N}_0)$ . Then for all  $\delta > 0$  there exists a constant  $c = c(\delta, r, q) > 0$  such that both the quantities*

$$\left\| \left( \left[ \sum_{k < j} 2^{-\delta(j-k)r} |x_k|^r \right]^{1/r} \right)_{j \in \mathbb{N}_0} \right\|_{\ell_q(\mathbb{N}_0)} \quad \text{and} \quad \left\| \left( \left[ \sum_{k \geq j} 2^{\delta(j-k)r} |x_k|^r \right]^{1/r} \right)_{j \in \mathbb{N}_0} \right\|_{\ell_q(\mathbb{N}_0)}$$

*are upper bounded by  $c \cdot \|(x_k)_{k \in \mathbb{N}_0}\|_{\ell_q(\mathbb{N}_0)}$ .*

The next assertion constitutes a generalization of Schur's Lemma to the case of  $\sigma$ -finite measure spaces. It can be shown by straightforward calculations along the lines of [22, (0.10)].

**Lemma 5.2.** *Let  $(X, \mu)$  and  $(Y, \nu)$  denote  $\sigma$ -finite measure spaces,  $1 \leq p \leq \infty$ , as well as  $1/p' + 1/p = 1$ . Moreover, assume that the measurable function  $K: X \times Y \rightarrow \mathbb{C}$  satisfies*

$$C_1(K) := \operatorname{esssup}_{x \in X} \int_Y |K(x, y)| \, d\nu(y) < \infty \quad \text{and} \quad C_2(K) := \operatorname{esssup}_{y \in Y} \int_X |K(x, y)| \, d\mu(x) < \infty.$$

*Then the integral operator  $T: L_p(X, \mu) \rightarrow L_p(Y, \nu)$ , given by*

$$f \mapsto Tf := \int_X K(x, \cdot) f(x) \, d\mu(x),$$

*is well-defined and satisfies  $\|T\| := \|T\|_{\mathcal{L}(L_p(X, \mu), L_p(Y, \nu))} \leq C_1(K)^{1/p} \cdot C_2(K)^{1/p'}$ .*

Finally, in the proof of Theorem 2.9 we make use of the following estimate. Therein  $\Gamma$  denotes an arbitrary set furnished with some metric.

**Lemma 5.3.** *Let  $\nabla = (\nabla_j)_{j \in \mathbb{N}_0}$  denote a multiscale grid of dimension  $d \in \mathbb{N}$  for some set  $\Gamma$  in the sense of Definition 2.1. Then, for each  $j, k \in \mathbb{N}_0$ , every  $s > d$  and all finite sets  $\mathcal{T} \neq \emptyset$ , we have*

$$\sup_{x \in \Gamma \times \mathcal{T}} \sum_{\xi \in \nabla_j} \frac{1}{[1 + 2^k \operatorname{dist}(\xi, x)]^s} \leq C \cdot \max\{1, 2^{(j-k)s}\}$$

*with some  $C > 0$  which does not depend on  $j$  and  $k$ .*

*Proof.* Let  $x \in \Gamma \times \mathcal{T}$  be fixed. Due to (A1) there exists  $\xi_x \in \nabla_j$  such that  $\operatorname{dist}(\xi_x, x) \leq c_1 2^{-j}$ . Hence, the triangle inequality implies  $\operatorname{dist}(\xi, x) \geq \operatorname{dist}(\xi, \xi_x) - c_1 2^{-j}$ , i.e.,

$$1 + 2^k \operatorname{dist}(\xi, x) \geq 1 - c_1 2^{k-j} + 2^k \operatorname{dist}(\xi, \xi_x) \quad \text{for all } \xi \in \nabla_j.$$

Next we define layer sets  $L := \{\xi \in \nabla_j \mid 0 \leq \text{dist}(\xi, \xi_x) < (\lceil c_1 \rceil + 1) 2^{-j}\}$ , and

$$L_\ell := \{\xi \in \nabla_j \mid \ell 2^{-j} \leq \text{dist}(\xi, \xi_x) < (\ell + 1) 2^{-j}\} \quad \text{for all } \ell \in \mathbb{N} \text{ with } \ell \geq \lceil c_1 \rceil + 1.$$

Since  $\nabla$  is uniformly well-separated (see (A2)), we note that  $\#L$  is bounded (uniformly in  $x$  and  $j$ ). The estimate  $\#L_\ell \lesssim \ell^{d-1}$  holds because of (A3). Moreover, we obviously have the coincidence  $\nabla_j = L \cup \bigcup_{\ell \geq \lceil c_1 \rceil + 1} L_\ell$ . Thus we may estimate

$$\begin{aligned} \sum_{\xi \in \nabla_j} \frac{1}{[1 + 2^k \text{dist}(\xi, x)]^s} &= \sum_{\xi \in L} \frac{1}{[1 + 2^k \text{dist}(\xi, x)]^s} + \sum_{\ell = \lceil c_1 \rceil + 1}^{\infty} \sum_{\xi \in L_\ell} \frac{1}{[1 + 2^k \text{dist}(\xi, x)]^s} \\ &\leq \#L + \sum_{\ell = \lceil c_1 \rceil + 1}^{\infty} \sum_{\xi \in L_\ell} \frac{1}{[1 - c_1 2^{k-j} + 2^k \text{dist}(\xi, \xi_x)]^s} \\ &\lesssim 1 + \sum_{\ell = \lceil c_1 \rceil + 1}^{\infty} \frac{\ell^{d-1}}{[1 + (\ell - \lceil c_1 \rceil) 2^{k-j}]^s} \\ &\leq 1 + 2^{(j-k)s} \sum_{n=1}^{\infty} \frac{(n + \lceil c_1 \rceil)^{d-1}}{n^s}, \end{aligned}$$

where the last sum converges due to the assumption  $s > d$ . Taking the supremum over all  $x \in \Gamma \times \mathcal{T}$  now completes the proof.  $\blacksquare$

## 5.2 Proofs

For the sake of completeness, in this final section, we add some proofs which were postponed in order to improve readability of the present manuscript. Let us start with showing the result on standard embeddings for sequence spaces  $b_{p,q}^\alpha(\nabla)$  stated in Proposition 2.5.

*Proof (of Proposition 2.5).* First of all we note that (A1)–(A3) in Definition 2.1 assure that every complex sequence  $\mathbf{a} := (a_{(j,\xi)})_{(j,\xi) \in \nabla}$  can be identified with some  $\tilde{\mathbf{a}} := (\tilde{a}_{(j,\lambda)})_{j \in \mathbb{N}_0, \lambda \in \tilde{\nabla}_j}$  such that for all parameters  $\alpha$ ,  $p$ , and  $q$  we have

$$\|\mathbf{a}\|_{b_{p,q}^\alpha(\nabla)} \sim \|\tilde{\mathbf{a}}\|_{b_{p,q}^\alpha(\tilde{\nabla})} \quad (20)$$

(with implied constants solely depending on  $p$  and  $q$ ), where the latter (quasi-)norm is defined by (2) with  $\nabla$  replaced by  $\tilde{\nabla} := (\tilde{\nabla}_j)_{j \in \mathbb{N}_0}$  and  $\tilde{\nabla}_j \subseteq \mathbb{Z}^d \times \{1, \dots, 2^d - 1\}$ . Moreover,  $\tilde{\nabla}$  satisfies (A4a) or (A4b), respectively, if and only if the same is true for  $\nabla$ .

In the case of finite index sets our definition of the spaces  $b_{p,q}^\alpha(\tilde{\nabla})$  exactly matches [11, Def. 3]. Then for  $\gamma \geq 0$  the claimed assertion is covered by [11, Lem. 4]. On the other

hand, if  $\gamma < 0$  then the sequence  $\tilde{\mathbf{a}}^* := (\tilde{a}_{(j,\lambda)}^*)_{j \in \mathbb{N}_0, \lambda \in \tilde{\nabla}_j}$  defined by  $\tilde{a}_{(j,\lambda)}^* := 2^{-j(d/2+\alpha+\gamma/2)}$  for every  $\lambda \in \tilde{\nabla}_j$ ,  $j \in \mathbb{N}_0$ , belongs to  $b_{p_0,q_0}^{\alpha+\gamma}(\tilde{\nabla}) \setminus b_{p_1,q_1}^\alpha(\tilde{\nabla})$ . Using (20) this contradicts  $b_{p_0,q_0}^{\alpha+\gamma}(\nabla) \hookrightarrow b_{p_1,q_1}^\alpha(\nabla)$  which completes the proof for this case.

Now let the index sets be infinite. Then, without loss of generality, we may assume that

$$\tilde{\nabla}_j = \begin{cases} \mathbb{Z}^d \times \{1\}, & j = 0, \\ \mathbb{Z}^d \times \{1, \dots, 2^d - 1\}, & j \in \mathbb{N}, \end{cases}$$

such that each sequence  $\tilde{\mathbf{a}}$  in the resulting spaces  $b_{p,q}^\alpha(\tilde{\nabla})$  can be identified with a function  $\tilde{f} \in B_{p,q}^\alpha(\mathbb{R}^d)$  by means of the isomorphism constructed in [33, Thm. 1.64]. Therein the Besov spaces  $B_{p,q}^\alpha(\mathbb{R}^d)$  are defined via harmonic analysis and the mapping  $\tilde{\mathbf{a}} \leftrightarrow \tilde{f}$  is accomplished by the use of Daubechies wavelets (which can be chosen as regular as we want in order to cover arbitrary parameter constellations). Moreover, we conclude

$$\|\mathbf{a}\|_{b_{p,q}^\alpha(\nabla)} \sim \|\tilde{\mathbf{a}}\|_{b_{p,q}^\alpha(\tilde{\nabla})} \sim \|\tilde{f}\|_{B_{p,q}^\alpha(\mathbb{R}^d)}$$

such that our claim in this case can be derived from corresponding embeddings of classical Besov (function) spaces  $B_{p,q}^\alpha(\mathbb{R}^d)$  which are well-known in the literature. Indeed, at the level of function spaces, sufficiency (and partially also necessity) of our conditions has been proven, e.g., in [25, Lem. 3]. To show the remaining “only if” parts we again construct counter examples at the level of sequence spaces: If  $p_0 > p_1$  then there certainly exists  $\mathbf{x} := (x_\lambda)_{\lambda \in \tilde{\nabla}_0} \in \ell_{p_0}(\tilde{\nabla}_0) \setminus \ell_{p_1}(\tilde{\nabla}_0)$ . Thus the sequence  $\tilde{\mathbf{a}}^*$  defined by  $\tilde{a}_{(0,\lambda)}^* := x_\lambda$ ,  $\lambda \in \tilde{\nabla}_0$ , and zero otherwise, belongs to  $b_{p_0,q_0}^{\alpha+\gamma}(\tilde{\nabla}) \setminus b_{p_1,q_1}^\alpha(\tilde{\nabla})$ . Hence,  $p_0 \leq p_1$  is necessary. It remains to check that for  $0 \leq \gamma < d(1/p_0 - 1/p_1)$  with  $p_0 \leq p_1$  the embedding is violated, too. For this purpose, we select one  $\lambda^*$  in each set  $\tilde{\nabla}_j$ ,  $j \in \mathbb{N}_0$ , and define  $\tilde{\mathbf{a}}^* := (\tilde{a}_{(j,\lambda)}^*)_{(j,\lambda) \in \tilde{\nabla}}$  by

$$\tilde{a}_{(j,\lambda)}^* := 2^{-j(\alpha+\gamma+d[1/2-1/p_0])}(1+j)^{-2/q_0} \quad \text{for } j \in \mathbb{N}_0 \quad \text{and } \lambda = \lambda^*,$$

and zero otherwise. Then it can be checked that again  $\tilde{\mathbf{a}}^* \in b_{p_0,q_0}^{\alpha+\gamma}(\tilde{\nabla}) \setminus b_{p_1,q_1}^\alpha(\tilde{\nabla})$  which completes the proof.  $\blacksquare$

It remains to deduce the properties of the three specific wavelet constructions  $\Psi_{\text{DS}}$ ,  $\Psi_{\text{HS}}$ , and  $\Psi_{\text{CTU}}$  stated in Lemma 3.4.

*Proof (of Lemma 3.4).* Since the support conditions (P1) and (P2) directly follow from the method which defines the wavelet systems, we are left with showing the cancellation-type property (P3). We split its proof into several steps according to the different wavelet constructions under consideration.

*Step 1.* First of all we deal with the case of (original) composite wavelets  $\Psi = \Psi_{\text{DS}}$  as constructed in [18] and follow the ideas indicated in [18, Sect. 4.7]. In order to improve transparency, for this step we stick to the (matrix-vector) notation used therein. In particular, by  $\Phi_j, \tilde{\Phi}_j$  we denote the vectors of primal and dual scaling functions at level  $j$  on  $\Gamma$ . For  $j \in \mathbb{N}_0$  we make use of the projectors

$$Q_j f := \langle f, \tilde{\Phi}_j \rangle \Phi_j \quad \text{and} \quad P_j^\Gamma f := \langle f, \tilde{\Lambda}_j \rangle \Phi_j,$$

defined in [18, Sect. 4.6], which map  $L_2(\Gamma)$  and  $\mathcal{C}(\Gamma)$  onto  $S(\Phi_j) \subset \mathcal{C}(\Gamma)$ , respectively. Here  $\tilde{\Lambda}_j$  denotes a vector of certain functionals such that  $P_j^\Gamma$  can be represented patchwise as

$$(P_j^\Gamma f)|_{\Gamma_i} = (P_j^\square(f \circ \kappa_i)) \circ \kappa_i^{-1}, \quad i = 1, \dots, N, \quad (21)$$

where  $P_j^\square$  are projectors acting on functions on the unit cube  $[0, 1]^d$ ; see Section 4.6 and 3.2 in [18] for details. Denoting the identity by  $\text{Id}$  we then have

$$Q_j P_j^\Gamma = P_j^\Gamma \quad \text{and} \quad (\text{Id} - Q_j)(\text{Id} - P_j^\Gamma) = \text{Id} - Q_j,$$

because of the duality of  $\Phi_j$  and  $\tilde{\Phi}_j$ . Clearly  $\tilde{\psi}_{j,\xi}^\Gamma \in S(\tilde{\Psi}_j^\Gamma) \perp S(\Phi_j) \ni Q_j(f)$  such that

$$\left| \langle f, \tilde{\psi}_{j,\xi}^\Gamma \rangle \right| = \left| \langle (\text{Id} - Q_j)f, \tilde{\psi}_{j,\xi}^\Gamma \rangle \right| = \left| \langle (\text{Id} - Q_j)(\text{Id} - P_j^\Gamma)f, \tilde{\psi}_{j,\xi}^\Gamma \rangle \right|,$$

where  $\tilde{\Psi}_j^\Gamma = (\tilde{\psi}_{j,\xi}^\Gamma)_{\xi \in \nabla_j^\Gamma}$  denotes the vector of dual wavelets at level  $j \in \mathbb{N}_0$  on  $\Gamma$ . Since the operators  $Q_j$  are uniformly bounded on  $L_2(\Gamma)$  (cf. [18, Rem. 4.6.1]) so is  $\text{Id} - Q_j$  and thus the normalization of  $\tilde{\psi}_{j,\xi}^\Gamma$  in  $L_2(\Gamma)$  gives

$$\begin{aligned} \left| \langle f, \tilde{\psi}_{j,\xi}^\Gamma \rangle \right| &\lesssim \left\| (\text{Id} - Q_j)(\text{Id} - P_j^\Gamma)f \right\|_{L_2(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma)} \left\| \tilde{\psi}_{j,\xi}^\Gamma \right\|_{L_2(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma)} \\ &\lesssim \left\| (\text{Id} - P_j^\Gamma)f \right\|_{L_2(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma)} \\ &\lesssim \sum_{i=1}^N \left\| (\text{Id} - P_j^\Gamma)f \right\|_{L_2(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma \cap \Gamma_i)}. \end{aligned}$$

Eq. (21), i.e., the patchwise representation of  $P_j^\Gamma$ , now implies that for every  $i = 1, \dots, N$

$$\begin{aligned} &\left\| (\text{Id} - P_j^\Gamma)f \right\|_{L_2(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma \cap \Gamma_i)} \\ &\leq \left\| f - p_i \circ \kappa_i^{-1} \right\|_{L_2(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma \cap \Gamma_i)} + \left\| (P_j^\square(f \circ \kappa_i) - p_i) \circ \kappa_i^{-1} \right\|_{L_2(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma \cap \Gamma_i)}, \end{aligned} \quad (22)$$



where the  $p_i$  denote arbitrarily chosen polynomials on  $[0, 1]^d$ .

Let us recall that, by construction, the operators  $P_j^\square$ , given by

$$g \mapsto P_j^\square g := \left\langle g, \tilde{\Lambda}_j^\square \right\rangle_{L_2([0,1]^d)} \Theta_j^\square,$$

are projections onto  $S(\Theta_j^\square)$ . Those spaces are strongly related to tensor products of shifts and dilates  $\theta_{j,k}$  of (boundary adapted) univariate  $D$ th-order cardinal B-splines  $\theta := {}_D\theta$ , where  $D := D^\Psi \in \mathbb{N}$ . Thus  $S(\Theta_j^\square)$  contains the space of all polynomials of total degree less than  $D^\Psi$  on the unit cube  $[0, 1]^d$ ; again see [18, Sect. 3.2]. The vectors  $\tilde{\Lambda}_j^\square$  consist of functionals  $\tilde{\lambda}_{j,k}^\square$  which are tensor products of  $L_2([0, 1])$ -inner products (with the duals  $\tilde{\theta}_{j,k}$  of  $\theta_{j,k}$ ) and (scaled) point evaluations at the boundary of the interval. Remember that it suffices to assume that  $g \in H^s([0, 1]^d)$  with  $s > d/2$  in order to ensure that sampling of the function  $g$  at points on the boundary of the unit cube is well-defined (in this case we find a continuous representer of  $g$  by means of Sobolev's embedding theorem).

Hence, if we restrict ourselves to polynomials  $p_i$  of degree smaller than  $D^\Psi$  then  $P_j^\square p_i = p_i$  and (22) can be rewritten as

$$\begin{aligned} & \left\| (\text{Id} - P_j^\Gamma) f \right\|_{L_2(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma \cap \Gamma_i)} \\ & \leq \left\| g_i \circ \kappa_i^{-1} \right\|_{L_2(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma \cap \Gamma_i)} + \left\| P_j^\square g_i \circ \kappa_i^{-1} \right\|_{L_2(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma \cap \Gamma_i)}, \end{aligned} \quad (23)$$

where we set  $g_i := f \circ \kappa_i - p_i$ . In order to bound the second summand in (23) we define the index sets

$$I_{j,\xi}^i := \left\{ k \mid \text{supp } \theta_{j,k}^\square \cap \kappa_i^{-1}(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma \cap \Gamma_i) \neq \emptyset \right\}, \quad i = 1, \dots, N,$$

for all tensor products  $\theta_{j,k}^\square$  in  $\Theta_j^\square$  whose support hit the set  $\kappa_i^{-1}(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma \cap \Gamma_i)$  in  $[0, 1]^d$ . Due to the local support of the  $\theta_{j,k}^\square$  the cardinality of this index sets is uniformly bounded in  $j$ ,  $\xi$ , and  $i$ . Therefore we conclude that

$$\begin{aligned} \left\| P_j^\square g_i \circ \kappa_i^{-1} \right\|_{L_2(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma \cap \Gamma_i)} &= \left\| P_j^\square g_i \right\|_{L_2(\kappa_i^{-1}(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma \cap \Gamma_i))} \\ &= \left\| \left\langle g_i, \tilde{\Lambda}_j^\square \right\rangle_{L_2([0,1]^d)} \Theta_j^\square \right\|_{L_2(\kappa_i^{-1}(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma \cap \Gamma_i))} \end{aligned}$$

can be estimated from above by

$$\left\| \sum_{k \in I_{j,\xi}^i} \left\langle g_i, \tilde{\lambda}_{j,k}^\square \right\rangle_{L_2([0,1]^d)} \theta_{j,k}^\square \right\|_{L_2([0,1]^d)} \lesssim \max_{k \in I_{j,\xi}^i} \left| \left\langle g_i, \tilde{\lambda}_{j,k}^\square \right\rangle_{L_2([0,1]^d)} \right| \left\| \theta_{j,k}^\square \right\|_{L_2([0,1]^d)}.$$

Using the normalization of  $\theta_{j,k}^\square$ , as well as the bound on  $\tilde{\lambda}_{j,k}^\square(g_i) := \langle g_i, \tilde{\lambda}_{j,k}^\square \rangle_{L_2([0,1]^d)}$  stated in [18, Ineq. (3.2.5)], we obtain

$$\begin{aligned} \left\| P_j^\square g_i \circ \kappa_i^{-1} \right\|_{L_2(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma \cap \Gamma_i)} &\lesssim 2^{-jd/2} \max_{k \in I_{j,\xi}^i} \left\| g_i \right\|_{L_\infty(\text{supp } \tilde{\theta}_{j,k}^\square)} \\ &\leq 2^{-jd/2} \left\| g_i \right\|_{L_\infty(\tilde{C}_{j,\xi}^i)}, \end{aligned}$$

where  $\tilde{C}_{j,\xi}^i$  denotes a cube in  $[0,1]^d$  that contains  $\kappa_i^{-1}(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma \cap \Gamma_i) \cup (\bigcup_{k \in I_{j,\xi}^i} \text{supp } \tilde{\theta}_{j,k}^\square)$ . Since, by construction, we have  $\text{diam}(\text{supp } \theta_{j,k}^\square \cup \text{supp } \tilde{\theta}_{j,k}^\square) \lesssim 2^{-j}$ , this cube can be chosen such that  $|\tilde{C}_{j,\xi}^i| \lesssim 2^{-jd}$ .

Moreover, our choice of  $\tilde{C}_{j,\xi}^i$  allows to bound the first summand in (23) by the same quantity:

$$\begin{aligned} \left\| g_i \circ \kappa_i^{-1} \right\|_{L_2(\text{supp } \tilde{\psi}_{j,\xi}^\Gamma \cap \Gamma_i)} &\leq \left\| g_i \right\|_{L_2(\tilde{C}_{j,\xi}^i)} \\ &\leq |\tilde{C}_{j,\xi}^i|^{1/2} \cdot \left\| g_i \right\|_{L_\infty(\tilde{C}_{j,\xi}^i)} \\ &\lesssim 2^{-jd/2} \left\| g_i \right\|_{L_\infty(\tilde{C}_{j,\xi}^i)}. \end{aligned} \tag{24}$$

In conclusion, for  $(f \circ \kappa_i) \in H^s(\tilde{C}_{j,\xi}^i)$  with  $s > d/2$ , we established the bound

$$\left| \langle f, \tilde{\psi}_{j,\xi} \rangle \right| \lesssim \sum_{i=1}^N 2^{-jd/2} \left\| g_i \right\|_{L_\infty(\tilde{C}_{j,\xi}^i)} = \sum_{i=1}^N 2^{-jd/2} \left\| f \circ \kappa_i - p_i \right\|_{L_\infty(\tilde{C}_{j,\xi}^i)}, \tag{25}$$

where the  $p_i$  are arbitrary polynomials of degree less than  $D^\Psi$  on  $[0,1]^d$ .

In order to show the desired estimate (15) we finally apply Whitney's bound which (adapted to our needs) takes the form

$$\inf_{p \in \Pi_{\lceil s \rceil}(\Omega)} \|F - p\|_{L_\infty(\Omega)} \lesssim |\Omega|^{s/d-1/2} |F|_{H^s(\Omega)}, \quad s \geq d/2.$$

Therein  $\Omega$  denotes some cube in  $\mathbb{R}^d$  with volume  $|\Omega|$ ,  $\Pi_{\lceil s \rceil}(\Omega)$  is the space of polynomials of degree less than  $\lceil s \rceil$  (the smallest integer larger or equal to  $s$ ), and  $|F|_{H^s(\Omega)}$  is the usual semi-norm of  $F$  in  $H^s(\Omega)$ . Accordingly,

$$\left| \langle f, \tilde{\psi}_{j,\xi} \rangle \right| \lesssim \sum_{i=1}^N 2^{-jd/2} 2^{-jd(s/d-1/2)} |f \circ \kappa_i|_{H^s(\tilde{C}_{j,\xi}^i)}, \quad d/2 < s \leq D^\Psi,$$

which proves the first assertion claimed in (P3) for  $\Psi = \Psi_{\text{DS}}$ .

The proof of the second part of (P3) is obtained by exactly the same arguments, where every primal quantity (denoted without tilde) is replaced by its dual analogue (with tilde) and vice versa.

*Step 2.* We turn to modified composite wavelets  $\Psi = \Psi_{\text{HS}}$  as constructed in [27]. In this case we can make use of the fact that the primal and the dual wavelets satisfy patchwise cancellation properties of order  $\tilde{D}^\Psi$  and  $D^\Psi$ , respectively. That is, we have

$$\langle \psi_{j,\xi}^\Gamma \circ \kappa_i, p_i \rangle_{L_2([0,1]^d)} = 0 \quad \text{and} \quad \langle \tilde{\psi}_{j,\xi}^\Gamma \circ \kappa_i, p_i \rangle_{L_2([0,1]^d)} = 0, \quad i = 1, \dots, N, \quad (26)$$

for all  $j$  and  $\xi$ , as well as every polynomial  $p_i$  of degree less than  $\tilde{D}^\Psi$  resp.  $D^\Psi$  on  $[0,1]^d$ . For the primal wavelets this has been shown in [27, Prop. 4.1], whereas the assertion at the dual side simply follows from biorthogonality of the dual wavelets with the primal scaling functions (which are exact of order  $D^\Psi$ ).

Now the derivation of the first part of (P3) for  $\Psi = \Psi_{\text{HS}}$  is straightforward. From the definition of the inner product  $\langle \cdot, \cdot \rangle$  and the patchwise cancellation property (26) we deduce

$$\left| \langle f, \tilde{\psi}_{j,\xi}^\Gamma \rangle \right| \leq \sum_{i=1}^N \left| \langle \tilde{\psi}_{j,\xi}^\Gamma \circ \kappa_i, f \circ \kappa_i \rangle_{L_2(\tilde{C}_{j,\xi}^i)} \right| = \sum_{i=1}^N \left| \langle \tilde{\psi}_{j,\xi}^\Gamma \circ \kappa_i, f \circ \kappa_i - p_i \rangle_{L_2(\tilde{C}_{j,\xi}^i)} \right|$$

for all  $p_i \in \Pi_{D^\Psi}(\tilde{C}_{j,\xi}^i)$ , where the cubes  $\tilde{C}_{j,\xi}^i \subset [0,1]^d$ ,  $i = 1, \dots, N$ , can be chosen according to the requirements in the statement of the lemma. Again we set  $g_i := f \circ \kappa_i - p_i$ , apply Cauchy-Schwarz, and use the  $L_2(\Gamma)$ -normalization of the dual wavelets to obtain

$$\left| \langle \tilde{\psi}_{j,\xi}^\Gamma \circ \kappa_i, f \circ \kappa_i - p_i \rangle_{L_2(\tilde{C}_{j,\xi}^i)} \right| \leq \|g_i\|_{L_2(\tilde{C}_{j,\xi}^i)} \lesssim 2^{-jd/2} \|g_i\|_{L_\infty(\tilde{C}_{j,\xi}^i)},$$

where the second estimate is derived as in (24). Hence, again we have shown (25) and (as in the previous step) the bound (15) is implied by Whitney's estimate.

Since the proof for the primal wavelets is obtained in the same way, we have shown the claim also for this case.

*Step 3.* The proof of (P3) for the construction  $\Psi = \Psi_{\text{CTU}}$  given in [2] can be performed literally as in the previous step. The needed patchwise cancellation property (26) for the primal wavelets can be found as Formula (3.12) in [3, Section 3.4.1].  $\blacksquare$

## Acknowledgements

The author likes to thank S. Dahlke and H. Harbrecht for several valuable discussions.

## References

- [1] J. Bergh and J. Löfström. *Interpolation Spaces. An Introduction*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin, 1976.
- [2] C. Canuto, A. Tabacco, and K. Urban. The wavelet element method part I: Construction and analysis. *Appl. Comput. Harmon. Anal.*, 6(1):1–52, 1999.
- [3] C. Canuto, A. Tabacco, and K. Urban. The wavelet element method part II: Realization and additional features in 2D and 3D. *Appl. Comput. Harmon. Anal.*, 8(2):123–165, 2000.
- [4] A. Cohen. *Numerical Analysis of Wavelet Methods*, volume 32 of *Studies in Mathematics and its Applications*. North-Holland, Amsterdam, 2003.
- [5] A. Cohen, W. Dahmen, and R. A. DeVore. Adaptive wavelet methods for elliptic operator equations: Convergence rates. *Math. Comp.*, 70:27–75, 2001.
- [6] A. Cohen, W. Dahmen, and R. A. DeVore. Adaptive wavelet methods II: Beyond the elliptic case. *J. Found. Comput. Math.*, 2(3):203–245, 2002.
- [7] A. Cohen and R. Masson. Wavelet adaptive method for second order elliptic problems: Boundary conditions and domain decomposition. *Numer. Math.*, 86(2):193–238, 2000.
- [8] S. Dahlke, W. Dahmen, and R. A. DeVore. Nonlinear approximation and adaptive techniques for solving elliptic operator equations. In W. Dahmen, A. Kurdila, and P. Oswald, editors, *Multiscale Wavelet Methods for Partial Differential Equations*, pages 237–283, San Diego, 1997. Academic Press.
- [9] S. Dahlke and R. A. DeVore. Besov regularity for elliptic boundary value problems. *Comm. Partial Differential Equations*, 22(1-2):1–16, 1997.
- [10] S. Dahlke, L. Diening, C. Hartmann, B. Scharf, and M. Weimar. Besov regularity of solutions to the  $p$ -Poisson equation. In preparation, 2014.
- [11] S. Dahlke, E. Novak, and W. Sickel. Optimal approximation of elliptic problems by linear and nonlinear mappings II. *J. Complexity*, 22(4):549–603, 2006.
- [12] S. Dahlke, T. Raasch, M. Werner, M. Fornasier, and R. Stevenson. Adaptive frame methods for elliptic operator equations: The steepest descent approach. *IMA J. Numer. Anal.*, 27(4): 717–740, 2007.
- [13] S. Dahlke and M. Weimar. Besov regularity for operator equations on patchwise smooth manifolds. Report 2013-03, Fachbereich Mathematik und Informatik, Philipps-Universität Marburg, 2013.

- [14] W. Dahmen. Wavelet and multiscale methods for operator equations. *Acta Numer.*, 6:55–228, 1997.
- [15] W. Dahmen, H. Harbrecht, and R. Schneider. Compression techniques for boundary integral equations. Asymptotically optimal complexity estimates. *SIAM J. Numer. Anal.*, 43(6):2251–2271, 2006.
- [16] W. Dahmen, H. Harbrecht, and R. Schneider. Adaptive methods for boundary integral equations: Complexity and convergence estimates. *Math. Comp.*, 76(259):1243–1274, 2007.
- [17] W. Dahmen and R. Schneider. Wavelets with complementary boundary conditions — Function spaces on the cube. *Result. Math.*, 34(3–4):255–293, 1998.
- [18] W. Dahmen and R. Schneider. Composite wavelet bases for operator equations. *Math. Comp.*, 68:1533–1567, 1999.
- [19] W. Dahmen and R. Schneider. Wavelets on manifolds I: Construction and domain decomposition. *SIAM J. Math. Anal.*, 31(1):184–230, 1999.
- [20] R. A. DeVore. Nonlinear approximation. *Acta Numer.*, 7:51–150, 1998.
- [21] R. A. DeVore, B. Jawerth, and V. Popov. Compression of wavelet decompositions. *Am. J. Math.*, 114(4):737–785, 1992.
- [22] G. B. Folland. *Introduction to Partial Differential Equations (2nd ed.)*. Princeton University Press, Princeton, NJ, 1995.
- [23] M. Frazier and B. Jawerth. A discrete transform and decomposition of distribution spaces. *J. Funct. Anal.*, 93(1):34–170, 1990.
- [24] M. Hansen.  $N$ -term approximation rates and Besov regularity for elliptic PDEs on polyhedral domains. Technical Report 2012-41, Seminar for Applied Mathematics, ETH Zürich, 2012. To appear in *J. Found. Comput. Math.*, 2014.
- [25] M. Hansen and W. Sickel. Best  $m$ -term approximation and Lizorkin-Triebel spaces. *J. Approx. Theory*, 163(8):923–954, 2011.
- [26] H. Harbrecht and R. Schneider. Biorthogonal wavelet bases for the boundary element method. *Math. Nachr.*, 269–270:167–188, 2004.
- [27] H. Harbrecht and R. Stevenson. Wavelets with patchwise cancellation properties. *Math. Comp.*, 75:1871–1889, 2006.

- [28] N. Kalton, S. Mayboroda, and M. Mitrea. Interpolation of Hardy-Sobolev-Besov-Triebel-Lizorkin spaces and applications to problems in partial differential equations. In L. De Carli and M. Milman, editors, *Interpolation Theory and Applications (Contemporary Mathematics 445)*, pages 121–177, Providence, RI, 2007. Amer. Math. Soc.
- [29] P. G. Lemarié. Bases d’ondelettes sur les groupes de lie satisfiés. *Bull. Soc. math. France*, 117:211–232, 1989.
- [30] M. Primbs. New stable biorthogonal spline-wavelets on the interval. *Result. Math.*, 57(1–2): 121–162, 2010.
- [31] T. Runst and W. Sickel. *Sobolev Spaces of Fractional Order, Nemytskij Operators and Non-linear Partial Differential Equations*. de Gruyter, Berlin, 1996.
- [32] V. S. Rychkov. On a theorem of Bui, Paluszyński, and Taibleson. In *Studies on the theory of differentiable functions of several variables and its applications. Part 18*, pages 280–292. Moskva: Nauka; Moskva: MAIK Nauka/Interperiodika, 1999.
- [33] H. Triebel. *Theory of Function Spaces III*. Birkhäuser, Basel, 2006.
- [34] H. Triebel. *Function Spaces and Wavelets on Domains*, volume 7 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2008.
- [35] H. Triebel. *Theory of Function Spaces*. Birkhäuser, Basel, reprint of the 1983 original edition, 2010.